GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:22:12 ; Search time 1446.77 Seconds
(without alignments)
11.932 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GGCCAAAAATTGAAAAACTA.....GCCCGGCCGCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28.6	41.4	2360	1	Q65321	Vaccinia virus 7.5
2	28.6	41.4	2360	1	Q86798	Vaccinia virus vec
3	28.6	41.4	4987	1	Q65322	Vaccinia virus pro
4	28.6	41.4	4987	1	Q86799	Vaccinia virus vec
5	27.8	40.3	207	1	Q40242	Vaccinia virus ear
6	26.6	38.6	38	1	Q74458	Poxvirus ATI promo
c 7	26.6	38.6	38	1	Q74459	Poxvirus ATI promo
8	26	37.7	870	1	T63623	DNA encoding L-pro
9	26	37.7	1081	1	T63622	KpnI-EcoRI fragmen
10	25.8	37.4	94	1	Q94699	Mycoplasma gallise
11	25.4	36.8	40	1	Q34776	MluI-BamHI fragmen
12	25.4	36.8	40	1	Q86231	Recombinant pox vi
13	25	36.2	40	1	Q68938	Vaccinia virus ear
14	25	36.2	40	1	Q70568	Sequence of early
15	25	36.2	40	1	T48506	Synthetic early pr
16	25	36.2	70	1	Q99648	Swinepox early/lat
17	25	36.2	102	1	Q99624	Homology vector 52
18	25	36.2	108	1	Q99635	Homology vector 53
19	25	36.2	178	1	Q68946	Junction B of SfiI
20	25	36.2	178	1	Q70553	Junction B of SfiI
21	25	36.2	178	1	T48503	Homology vector 50
22	25	36.2	182	1	Q68945	Junction A of SfiI
23	25	36.2	182	1	Q70552	Junction A of SfiI
24	25	36.2	182	1	T48502	Homology vector 50
25	25	36.2	205	1	Q99629	Homology vector 53
26	25	36.2	352	1	Q46676	Sequence Inserted
27	25	36.2	438	1	Q46679	Sequence Inserted
28	25	36.2	523	1	Q46677	Sequence Inserted
29	25	36.2	988	1	T33620	DNA encoding SecA
30	25	36.2	4177	1	Q68943	SfiI fragment cont
31	25	36.2	4177	1	Q70570	SfiI fragment cont
32	25	36.2	4177	1	T48510	SfiI fragment enco
33	24.8	35.9	226	1	T43285	Promoter for infec
34	24.8	35.9	1433	1	X38293	M. tuberculosis se
35	24.8	35.9	3189	1	T31315	Adenovirus E4 and
36	24.8	35.9	8710	1	V32370	Complete sequence
37	24.8	35.9	10610	1	V32375	Complete sequence
38	24.8	35.9	32026	1	T60559	Recombinant adenov
39	24.8	35.9	34303	1	V07261	Adenoviral vector
40	24.8	35.9	34382	1	X15627	Recombinant adenov
41	24.8	35.9	34427	1	X07371	Adenovirus vector
42	24.8	35.9	35000	1	T60557	Recombinant cis-ac
43	24.8	35.9	35935	1	V07258	Adenovirus 5 genom
44	24.8	35.9	36538	1	T60558	Recombinant trans-
45	24.6	35.7	1717	1	V00493	Human tumour necro

ALIGNMENTS

RESULT 1
Q65321
ID Q65321 standard; DNA; 2360 BP.
AC Q65321;
DT 21-DEC-1994 (first entry)
DE Vaccinia virus 7.5k protein promoter PCR fragment.
KW Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;
KW antisense; therapy; inhibition; viral protein precursor;
KW recombinant vaccinia virus; 7.5k protein promoter;
KW HCV core protein; luciferase; fusion construct; ds.
OS Vaccinia virus
PN CA2104649-A.
PD 26-FEB-1994.
PF 23-AUG-1993; 104649.
PR 25-AUG-1992; JP-248796.
PR 03-MAR-1993; JP-042736.
PA (SEKI/) SEKI M.
PI Honda Y, Seki M, Yamada E;
DR WPI; 94-151836/19.
PT Anti:sense oligo:nucleotide(s) complementary to the hepatitis C
PT virus genome - are useful as antiviral agents
PS Example 5; Page 241-242; 262pp; English.
CC Synthetic primers Q65314 and Q65315 or Q65316 and Q65317 were
CC designed for amplification and modification of the 7.5k protein
CC promoter of vaccinia virus strain WR. A plasmid designated pHASE
CC contained the vaccinia viral promoter inserted in the same
CC direction as the HA gene. The sequence of the SalI-HindIII fragment
CC from pHASE was determined (Q65321). This promoter fragment was
CC used in the construction of recombinant vaccinia virus pVV5CL which
CC codes for a HCV core protein-luciferase fusion protein. The
CC construct was useful for assaying the inhibitory activity of
CC various antisense oligonucleotides on HCV gene translation.
SQ Sequence 2360 BP; 857 A; 410 C; 400 G; 693 T;

Query Match 41.4%; Score 28.6; DB 1; Length 2360;
Best Local Similarity 88.6%; Pred. No. 0.25;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
||||||||||||||||||| ||||||||||||
Db 1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269

RESULT 2
Q86798
ID Q86798 standard; DNA; 2360 BP.
AC Q86798;
DT 04-DEC-1995 (first entry)
DE Vaccinia virus vector pHASE.
KW Vaccinia virus vector; pHASE; hepatitis C virus; T7N1-19;
KW antiviral agent; poliovirus; human rhinovirus;
KW internal ribozyme entry site; non-A non-B;
KW cerebral cardio-hepatitis; foot and mouth disease; ds.
OS Vaccinia virus.
PN J07069899-A.

```
PD   14-MAR-1995.
PF   02-SEP-1993; 241973.
PR   02-SEP-1993; JP-241973.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-144713/19.
PT   Antiviral agent comprising component which disrupts viral gene
PT   translation - used for the selective inhibition of e.g. Hepatitis
PT   C virus, polio:virus and human rhinovirus
PS   Disclosure; Page 18; 23pp; Japanese.
CC   Q86788 encodes R72800 Hepatitis C virus (HCV) T7N1-19, which
CC   disrupts viral gene translation, by preventing the binding of
CC   the viral mRNA to the internal ribozyme entry site. It was
CC   used in the construction of an antiviral agent detecting vector,
CC   which comprises a vaccinia virus vector (Q86798), a vector
CC   containing T7N1-19 and a firefly luciferase gene. The antiviral
CC   agent can be used to treat HCV, poliovirus, cerebral cardio-
CC   hepatitis, human rhinovirus and foot and mouth disease viral
CC   infections.
SQ   Sequence   2360 BP;    858 A;    410 C;    399 G;    693 T;

  Query Match             41.4%;  Score 28.6;  DB 1;  Length 2360;
  Best Local Similarity   88.6%;  Pred. No. 0.25;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
           ||||||||||||||||||    ||||||||||||
Db    1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269


RESULT   3
Q65322
ID   Q65322 standard; DNA; 4987 BP.
AC   Q65322;
DT   21-DEC-1994  (first entry)
DE   Vaccinia virus promoter and HCV core protein-luciferase fusion gene.
KW   Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;
KW   antisense; therapy; inhibition; viral protein precursor;
KW   recombinant vaccinia virus; HCV core protein gene; firefly;
KW   luciferase reporter gene; fusion construct; ds.
OS   Chimeric Vaccinia virus.
OS   Chimeric Hepatitis C virus.
OS   Chimeric Photinus pyralis.
FH   Key             Location/Qualifiers
FT   cds             1826. .4057
FT                   /*tag= a
FT                   /note= "encodes HCV core protein-luciferase
FT                   fusion protein"
PN   CA2104649-A.
PD   26-FEB-1994.
PF   23-AUG-1993; 104649.
PR   25-AUG-1992; JP-248796.
PR   03-MAR-1993; JP-042736.
PA   (SEKI/) SEKI M.
PI   Honda Y,  Seki M,  Yamada E;
DR   WPI; 94-151836/19.
DR   P-PSDB; R54867.
PT   Anti:sense oligo:nucleotide(s) complementary to the hepatitis C
PT   virus genome - are useful as antiviral agents
PS   Example 5; Page 243-251; 262pp; English.
CC   A recombinant vaccinia virus which codes for a HCV core protein
CC   fused to the firefly luciferase enzyme was constructed from PCR
CC   amplified fragments. The construct was useful for assaying the
CC   inhibitory activity of various antisense oligonucleotides on HCV
CC   gene translation.
SQ   Sequence   4987 BP;    1520 A;    1052 C;    1083 G;    1332 T;

  Query Match             41.4%;  Score 28.6;  DB 1;  Length 4987;
  Best Local Similarity   88.6%;  Pred. No. 0.31;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
           ||||||||||||||||||    ||||||||||||
Db    1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269


RESULT   4
Q86799
ID   Q86799 standard; DNA; 4987 BP.
AC   Q86799;
DT   04-DEC-1995  (first entry)
DE   Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.
KW   Vaccina virus vector; firefly luciferase; hepatitis C virus; T7N1-19;
KW   antiviral agent; poliovirus; human rhinovirus;
KW   internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;
KW   foot and mouth disease; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   cds             1826. .4057
FT                   /*tag= a
PN   J07069899-A.
PD   14-MAR-1995.
PF   02-SEP-1993; 241973.
PR   02-SEP-1993; JP-241973.
PA   (MITU ) MITSUBISHI KASEI CORP.
DR   WPI; 95-144713/19.
DR   P-PSDB; R72801.
PT   Antiviral agent comprising component which disrupts viral gene
PT   translation - used for the selective inhibition of e.g. Hepatitis
PT   C virus, polio:virus and human rhinovirus
PS   Disclosure; Pages 19-22; 23pp; Japanese.
CC   Q86788 encodes R72800 Hepatitis C virus (HCV) T7N1-19, which
CC   disrupts viral gene translation, by preventing the binding of
CC   the viral mRNA to the internal ribozyme entry site. It was
CC   used in the construction of an antiviral agent detecting vector
CC   (Q86799 which encodes R72801), which comprises a vaccinia virus
CC   vector, a vector containing T7N1-19 and a firefly luciferase gene.
CC   The antiviral agent can be used to treat HCV, poliovirus, cerebral
CC   cardio-hepatitis, human rhinovirus and foot and mouth disease viral
CC   infections.
SQ   Sequence   4987 BP;    1516 A;    1050 C;    1081 G;    1322 T;

  Query Match             41.4%;  Score 28.6;  DB 1;  Length 4987;
  Best Local Similarity   88.6%;  Pred. No. 0.31;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
           ||||||||||||||||||    ||||||||||||
Db    1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269


RESULT   5
Q40242
ID   Q40242 standard; DNA; 207 BP.
AC   Q40242;
DT   10-AUG-1993  (first entry)
DE   Vaccinia virus early promoter.
KW   HCV; hepatitis C virus; nonB-nonB hepatitis; NANBHV; vaccinia virus;
KW   VV; cowpox; liver disease; infection; diagnosis; testing; antigen;
KW   immunoassay.
OS   Synthetic.
PN   J05078395-A.
PD   30-MAR-1993.
PF   20-SEP-1991; 241275.
PR   20-SEP-1991; JP-241275.
PA   (JAPG ) NIPPON ZEON KK.
DR   WPI; 93-140360/17.
DR   P-PSDB; R34635.
PT   Glyco-protein derived from blood disseminating non A non B
PT   hepatitis virus - used as diagnostic agent for vaccinia virus
PT   infection
PS   Disclosure; Page 9; 11pp; Japanese.
CC   This sequence is the vaccinia virus early promoter, used when
CC   cloning a glycoprotein originating from blood disseminating nonA-nonB
```

```
CC   hepatitis virus (NANBHV) and causing an antigen-antibody reaction with
CC   anti-vaccinia virus. NANBHV originated core and envelope genes were
CC   cloned. The resultant 4.01kb plasmid was named pIK4CE and its sequence
CC   determined. A recombinant vector contg. vaccinia virus (VV) TK gene with
CC   with the VV 7.5kb promoter was prepd. This vector contd. VV originating
CC   TK gene with the synthetic VV early promoter. A second recombinant vector
CC   for the prepn. of vv. having the structural protein gene region of NANBHV
CC   was prepd, and used to produce a recombinant VV. The NANBHV protein was
CC   expressed in a VV infected cell.
SQ   Sequence   207 BP;    66 A;    48 C;    38 G;    55 T;


  Query Match            40.3%;  Score 27.8;  DB 1;  Length 207;
  Best Local Similarity  74.5%;  Pred. No. 0.25;
  Matches   35;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGC 51
          |||||||||||||||||     ||||||||||   | |  ||||||
Db     79 AAAAATTGAAAAACTATTCTAATTTATTGCAATAGTCGACCATGGCC 125


RESULT   6
Q74458
ID   Q74458 standard; DNA; 38 BP.
AC   Q74458;
DT   22-MAY-1995  (first entry)
DE   Poxvirus ATI promoter, nucleotides -28 to +10.
KW   ATI promoter; poxvirus; pro-phase expression; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    35. .38
FT                   /*tag= a
FT                   /label= sticky_end
FT                   /note= "see Q74459 for complementary strand"
PN   J06237773-A.
PD   30-AUG-1994.
PF   21-JAN-1993; 008424.
PR   22-DEC-1992; JP-342362.
PA   (SHID/) SHIDA H.
PA   (TOFU ) TONEN CORP.
DR   WPI; 94-312814/39.
PT   A poxvirus promoter contg. an ATI promoter and prophase - useful
PT   for the production of exotic genes in large quantities
PS   Claim 2; Page 2; 10pp; Japanese.
CC   Q74458 and Q74459 together form a double-stranded nucleotide
CC   fragment with overhanging sticky ends. The fragment comprises the
CC   ATI promoter and pro-phase expression promoter. Coding sequences
CC   (e.g. heterologous coding sequences or homologous sequences such as
CC   the poxvirus haemagglutinin gene) can be placed downstream of the
CC   promoter fragment for recombinant production of the polypeptides.
SQ   Sequence   38 BP;    15 A;    4 C;    7 G;    12 T;


  Query Match            38.6%;  Score 26.6;  DB 1;  Length 38;
  Best Local Similarity  87.9%;  Pred. No. 0.41;
  Matches   29;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||    ||||||||||||
Db      1 AAAAATTGAAAAACTAGTCTAATTTATTGCACG 33


RESULT   7
Q74459/c
ID   Q74459 standard; DNA; 38 BP.
AC   Q74459;
DT   22-MAY-1995  (first entry)
DE   Poxvirus ATI promoter, nucleotides -32 to +6 (complementary strand).
KW   ATI promoter; poxvirus; pro-phase expression; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    35. .38
FT                   /*tag= a
FT                   /label= sticky_end
FT                   /note= "see Q74458 for complementary strand"
PN   J06237773-A.
PD   30-AUG-1994.
PF   21-JAN-1993; 008424.
PR   22-DEC-1992; JP-342362.
PA   (SHID/) SHIDA H.
PA   (TOFU ) TONEN CORP.
DR   WPI; 94-312814/39.
PT   A poxvirus promoter contg. an ATI promoter and prophase - useful
PT   for the production of exotic genes in large quantities
PS   Claim 2; Page 2; 10pp; Japanese.
CC   Q74458 and Q74459 together form a double-stranded nucleotide
CC   fragment with overhanging sticky ends. The fragment comprises the
CC   ATI promoter and pro-phase expression promoter. Coding sequences
CC   (e.g. heterologous coding sequences or homologous sequences such as
CC   the poxvirus haemagglutinin gene) can be placed downstream of the
CC   promoter fragment for recombinant production of the polypeptides.
SQ   Sequence   38 BP;    12 A;    7 C;    4 G;    15 T;


  Query Match            38.6%;  Score 26.6;  DB 1;  Length 38;
  Best Local Similarity  87.9%;  Pred. No. 0.41;
  Matches   29;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||    ||||||||||||
Db     34 AAAAATTGAAAAACTAGTCTAATTTATTGCACG 2


RESULT   8
T63623
ID   T63623 standard; DNA; 870 BP.
AC   T63623;
DT   31-MAY-1997  (first entry)
DE   DNA encoding L-proline-3-hydroxylase.
KW   L-proline-3-hydroxylase; L-proline; cis-3-hydroxy-L-proline;
KW   2-ketoglutaric acid; ferrous ion; intermediate; drug; food additive; ds.
OS   Streptomyces sp. strain TH1.
PN   WO9627668-A1.
PD   12-SEP-1996.
PF   07-MAR-1996; J00558.
PR   07-MAR-1995; JP-046987.
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PI   Mori H,  Ozaki A,  Shibasaki T;
DR   WPI; 96-425428/42.
DR   P-PSDB; W09782.
PT   DNA coding for L-proline-3-hydroxylase of microbial origin - for
PT   large scale production of 3-hydroxy-L-proline useful as drug
PT   intermediate or food additive
PS   Claim 3; Page 48; 81pp; Japanese.
CC   The present sequence, derived from pTH30, isolated from Streptomyces
CC   sp., encodes an L-proline-3-hydroxylase (W09782) which converts
CC   L-proline to cis-3-hydroxy-L-proline in the presence of 2-ketoglutaric
CC   acid and ferrous ions. The DNA is used for efficient production of
CC   cis-3-hydroxy-L-proline on an industrial scale for use as an
CC   intermediate in the manufacture of drugs and as a food additive.
SQ   Sequence   870 BP;    173 A;    299 C;    262 G;    136 T;


  Query Match            37.7%;  Score 26;  DB 1;  Length 870;
  Best Local Similarity  70.0%;  Pred. No. 1.6;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
          | ||||||  ||  |  ||||||||||| |   | |||||| ||||  ||
Db    441 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 490


RESULT   9
T63622
ID   T63622 standard; DNA; 1081 BP.
AC   T63622;
```

```
DT   03-JUN-1997  (first entry)
DE   KpnI-EcoRI fragment of pTH30 encoding L-proline-3-hydroxylase.
KW   L-proline-3-hydroxylase; L-proline; cis-3-hydroxy-L-proline;
KW   2-ketoglutaric acid; ferrous ion; intermediate; drug; food additive; ds.
OS   Streptomyces sp. strain TH1.
FH   Key             Location/Qualifiers
FT   cds             51. .923
FT                   /*tag= a
FT                   /note= "encodes W09782"
PN   WO9627668-A1.
PD   12-SEP-1996.
PF   07-MAR-1996; J00558.
PR   07-MAR-1995; JP-046987.
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PI   Mori H,  Ozaki A,  Shibasaki T;
DR   WPI; 96-425428/42.
DR   P-PSDB; W09782.
PT   DNA coding for L-proline-3-hydroxylase of microbial origin - for
PT   large scale production of 3-hydroxy-L-proline useful as drug
PT   intermediate or food additive
PS   Claim 3; Page 53-54; 81pp; Japanese.
CC   This KpnI-EcoRI fragment, derived from pTH30, which was isolated from
CC   Streptomyces sp., contains the open reading frame (T63623) that encodes
CC   an L-proline-3-hydroxylase (W09782) which converts L-proline to cis-3-
CC   hydroxy-L-proline in the presence of 2-ketoglutaric acid and ferrous
CC   ions. The DNA is used for efficient production of cis-3-hydroxy-L-proline
CC   on an industrial scale for use as an intermediate in the manufacture of
CC   drugs and as a food additive.
SQ   Sequence   1081 BP;    218 A;    365 C;    334 G;    164 T;

  Query Match            37.7%;  Score 26;  DB 1;  Length 1081;
  Best Local Similarity  70.0%;  Pred. No. 1.7;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
          | ||||||  ||  |  |||||||||||  |   | ||||||  ||||  ||
Db    491 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 540


RESULT  10
Q94699
ID   Q94699 standard; DNA; 94 BP.
AC   Q94699;
DT   13-MAR-1996  (first entry)
DE   Mycoplasma gallisepticum antigenic protein TM-16 gene promoter.
KW   Antigenic protein; vaccine; poultry; diagnosis; TM-16;
KW   promoter; ds.
OS   Synthetic.
PN   J07133295-A.
PD   23-MAY-1995.
PF   27-AUG-1993; 213102.
PR   27-AUG-1993; JP-213102.
PA   (SHIO ) SHIONOGI & CO LTD.
DR   WPI; 95-220782/29.
PT   A new antigenic protein which reacts with Mycoplasma gallisepticum -
PT   is useful in a component vaccine for use against poultry infected
PT   with M. gallisepticum.
PS   Example 1; Page 8; 33pp; Japanese.
CC   Q94699 is a promoter for the Mycoplasma gallisepticum antigenic
CC   protein TM-16 gene. TM-16 can be used as a vaccine for M.
CC   gallisepticum infectious diseases in poultry, and as a diagnostic
CC   agent for M. gallisepticum.
SQ   Sequence   94 BP;    32 A;    11 C;    8 G;    43 T;

  Query Match            37.4%;  Score 25.8;  DB 1;  Length 94;
  Best Local Similarity  81.1%;  Pred. No. 1;
  Matches   30;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGC 41
          |||||||||||||||     |||||||||| || |
Db     58 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGTC 94


RESULT  11
Q34776
ID   Q34776 standard; DNA; 40 BP.
AC   Q34776;
DT   07-MAY-1993  (first entry)
DE   MluI-BamHI fragment of pNZ1729R insertion vector.
KW   Insertion vector; fowlpox virus; FPV; strain NP; MDV;
KW   glycoprotein B homologue; ds.
OS   Synthetic.
PN   EP-520753-A.
PD   30-DEC-1992.
PF   24-JUN-1992; 305775.
PR   28-JUN-1991; US-722860.
PR   10-DEC-1991; US-803633.
PA   (JAPG ) NIPPON ZEON KK.
PA   (USDA ) US SEC OF AGRIC.
PI   Lee LF, Li Y, Nazerian K, Ogawa R, Yanagida N;
DR   WPI; 93-001546/01.
PT   Recombinant fowl pox virus contg. Marek's disease virus antigen
PT   gene - used to produce cell-free vaccine against Marek's disease
PT   virus
PS   Example 1; Fig 2; 30pp; English.
CC   The 5'end of this strand overhangs the 3'end of its complementary
CC   strand by 4 bases (i.e. 5'-CGCG-3'); the 5' end of the complementary
CC   strand overhangs the 3'end of this strand, also by 4 bases (i.e.
CC   5'-GATC-3'). The double-stranded fragment was ligated to other
CC   synthetic fragments containing a multiple cloning site (mcs) and a
CC   2 directional poxvirus early transcriptional termination signal.
CC   These were ligated downstream of a FPV promoter, initiation codon
CC   and lacZ gene (from pMC1871 and pMA001). The resulting 3.5kb
CC   EcoRI-HindIII fragment was inserted in pNZ173SR to make the pNZ1729R
CC   insertion vector. (Plasmid pNZ173R was made by subcloning a 3kb HpaI-
CC   SpeI fragment from a 7.3kb EcoRI fragment of FPV NP strain in pUC18.
CC   All mcs's were eliminated from both junction regions and an mcs was
CC   inserted into 2 adjacent EcoRV sites in the cloned FPV fragment
CC   with linkers). A fragment contg. the entire coding region of Marek
CC   Disease Virus glycoprotein B homologue (gBh) was inserted into
CC   pNZ1729R to produce a recombinant FPV/MDVgBh virus for immunising
CC   chickens.  See also Q34774-Q34781.
SQ   Sequence   40 BP;    15 A;    7 C;    5 G;    13 T;

  Query Match            36.8%;  Score 25.4;  DB 1;  Length 40;
  Best Local Similarity  82.9%;  Pred. No. 1.1;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||     |||||||||| ||
Db      6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


RESULT  12
Q86231
ID   Q86231 standard; DNA; 40 BP.
AC   Q86231;
DT   10-JAN-1996  (first entry)
DE   Recombinant pox virus E promoter fragment.
KW   promoter; pox virus; combined promoter; recombinant; vaccine;
KW   poultry; live; ds.
OS   Synthetic.
PN   J07067655-A.
PD   14-MAR-1995.
PF   30-AUG-1993; 238953.
PR   30-AUG-1993; JP-238953.
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
DR   WPI; 95-143849/19.
PT   Combined pox:virus promoters contain at least four promoter
PT   fragments  - and pox:virus containing the combined promoters and a
PT   foreign gene
PS   Claim 4; Fig 1; 12pp; Japanese.
```

```
CC   The DNA shown is a pref. DNA fragment with E (early) promoter activity,
CC   which is used in the construction of a combined promoter for recombinant
CC   poxvirus. Combined promoters for poxvirus comprise at least four DNA
CC   fragments having a promoter activity, pref. these fragments have early
CC   and late activity, and the promoter has at least one, pref. more than two
CC   promoters.  The combined promoter has a strong activity and hence is
CC   useful in an effective vaccine, esp. against poxvirus in poultry.
SQ   Sequence   40 BP;    15 A;    7 C;    5 G;    13 T;


  Query Match            36.8%;  Score 25.4;  DB 1;  Length 40;
  Best Local Similarity  82.9%;  Pred. No. 1.1;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||     |||||||||||  ||
Db      6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


RESULT  13
Q68938
ID   Q68938 standard; DNA; 40 BP.
AC   Q68938;
DT   13-APR-1995  (first entry)
DE   Vaccinia virus early promoter cassette EP1.
KW   Early promoter; cassetter; recombinant fowlpox vector;
KW   synthetic pox promoter; ss.
OS   Synthetic.
PN   WO9419014-A.
PD   01-SEP-1994.
PF   28-FEB-1994; U01826.
PR   26-FEB-1993; US-024156.
PA   (JAPG ) NIPPON ZEON KK.
PA   (SYTR ) SYNTRO CORP.
PI   Cochran MD;
DR   WPI; 94-294007/36.
PT   New recombinant fowl pox virus for use in vaccines - contains
PT   genes expressing antigens of Newcastle disease virus and opt.
PT   infectious bronchitis virus
PS   Disclosure; Page 62; 85pp; English.
CC   Synthetic pox viral promoters were constructed for recombinant
CC   fowlpox vectors. The four promoter cassettes - EP1, LP1, EP2 and LP2
CC   are based on promoters that have been defined in the Vaccinia virus.
CC   Each cassette was designed to contain DNA sequences defined in
CC   Vaccinia flanked by restriction sites. Initiator Mets were also
CC   designed into each cassette such that in-frame fusions could be made
CC   at either EcoRI or BamHI sites. A set of translational stop codons
CC   in all three reading frames and an early transcriptional termination
CC   signal was also engineered downstream of the in-frame fusion site.
SQ   Sequence   40 BP;    17 A;    5 C;    5 G;    13 T;


  Query Match            36.2%;  Score 25;  DB 1;  Length 40;
  Best Local Similarity  84.8%;  Pred. No. 1.6;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||     ||||||||||||
Db      1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


RESULT  14
Q70568
ID   Q70568 standard; DNA; 40 BP.
AC   Q70568;
DT   14-APR-1995  (first entry)
DE   Sequence of early promoter cassette EP1.
KW   Early promoter; synthetic pox viral promoter; fowlpox; vaccinia; ss.
OS   Synthetic.
PN   WO9419015-A.
PD   01-SEP-1994.
PF   28-FEB-1994; U02252.
PR   26-FEB-1993; US-024156.
PA   (SYTR ) SYNTRO CORP.
PI   Cochran MD;
DR   WPI; 94-294008/36.
PT   New recombinant fowl pox viruses - useful as vaccines against
PT   fowl pox virus, Newcastle Disease Virus and infectious
PT   laryngotracheitis virus.
PS   Disclosure; Page 71; 97pp; English.
CC   Four promoter cassettes were designed based on promoters that have
CC   been defined in vaccinia virus. These are EP1, LP1, EP2 and LP2
CC   (Q70568, Q70571, Q70572 and Q70569 respectively). Each cassette was
CC   designed to contain the DNA sequences defined in vaccinia flanked by
CC   restriction sites which could be used to combine the cassettes in
CC   any order or combination. Initiator Mets were also designed into
CC   each cassette such that inframe fusions could be made at either
CC   EcoRI or BamHI sites. A set of translational stop codons in all
CC   three reading frames and an early transcriptional termination signal
CC   was also engineered downstream of the inframe fusion site.
SQ   Sequence   40 BP;    17 A;    5 C;    5 G;    13 T;


  Query Match            36.2%;  Score 25;  DB 1;  Length 40;
  Best Local Similarity  84.8%;  Pred. No. 1.6;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||     ||||||||||||
Db      1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


RESULT  15
T48506
ID   T48506 standard; DNA; 40 BP.
AC   T48506;
DT   05-MAY-1997  (first entry)
DE   Synthetic early promoter 1 (EP1).
KW   Fowlpox virus; FPV; recombinant virus; vector; vaccine; poultry;
KW   immunisation; Newcastle disease virus; NDV; promoter; EP1; ds.
OS   Synthetic.
PN   WO9640880-A1.
PD   19-DEC-1996.
PF   04-JUN-1996; U11187.
PR   07-JUN-1995; US-484790.
PA   (SYTR ) SYNTRO CORP.
PI   Cochran MD,  Junker DE,  Singer PA;
DR   WPI; 97-087060/08.
PT   New recombinant fowlpox virus - contg. a foreign DNA sequence
PT   inserted into the fowlpox virus genome, used for the prodn. of
PT   vaccines.
PS   Disclosure; Page 101; 134pp; English.
CC   4 Promoter cassettes (T48506-09) are based on promoters that have
CC   been defined in vaccinia virus.  They respectively comprise early
CC   promoter 1 (EP1), late promoter 1 (LP1), early promoter 2 (EP2)
CC   and late promoter 2 (LP2).  The promoters offer several advantages
CC   including the ability to control the strength and timing of foreign
CC   gene expression, and have been utilised in recombinant fowlpox
CC   virus vectors (see also T48502-03) to control expression of foreign
CC   DNA.  Recombinant fowlpox viruses can be used as vectors for the
CC   delivery of specific vaccine antigens or cytokines to animals, esp.
CC   poultry.
SQ   Sequence   40 BP;    17 A;    5 C;    5 G;    13 T;


  Query Match            36.2%;  Score 25;  DB 1;  Length 40;
  Best Local Similarity  84.8%;  Pred. No. 1.6;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||     ||||||||||||
Db      1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


Search completed: May 29, 2000, 21:58:06
```

Job time: 38154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:23:19 ; Search time 621.83 Seconds
(without alignments)
14.424 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GGCCAAAAATTGAAAAACTA.....GCCCGGCCGCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	26.4	38.3	40	1	US-07-803-633A-6	Sequence 6, Appli
c	2	26.4	38.3	40	3	US-08-525-742-16	Sequence 16, Appl
c	3	26.4	38.3	96	3	US-08-525-742-30	Sequence 30, Appl
	4	26	37.7	870	4	US-08-708-856A-3	Sequence 3, Appli
	5	26	37.7	870	5	US-09-287-375-3	Sequence 3, Appli
	6	26	37.7	1081	4	US-08-708-856A-13	Sequence 13, Appl
	7	26	37.7	1081	5	US-09-287-375-13	Sequence 13, Appl
c	8	25.8	37.4	93	3	US-08-525-742-28	Sequence 28, Appl
	9	25.8	37.4	97	3	US-08-525-742-27	Sequence 27, Appl
	10	25.4	36.8	40	1	US-07-803-633A-5	Sequence 5, Appli
	11	25.4	36.8	40	3	US-08-525-742-15	Sequence 15, Appl
	12	25.4	36.8	95	3	US-08-525-742-29	Sequence 29, Appl
	13	25	36.2	40	4	US-08-484-575A-8	Sequence 8, Appli
	14	25	36.2	40	5	US-08-477-459-8	Sequence 8, Appli
	15	25	36.2	40	6	PCT-US94-01826A-8	Sequence 8, Appli
	16	25	36.2	40	6	PCT-US94-02252A-8	Sequence 8, Appli
	17	25	36.2	70	3	US-08-097-554A-43	Sequence 43, Appl
	18	25	36.2	70	5	US-08-480-640A-43	Sequence 43, Appl
	19	25	36.2	102	1	US-07-820-154A-10	Sequence 10, Appl
	20	25	36.2	102	3	US-08-097-554A-10	Sequence 10, Appl
	21	25	36.2	102	5	US-08-480-640A-10	Sequence 10, Appl
	22	25	36.2	102	6	PCT-US93-00324-10	Sequence 10, Appl
	23	25	36.2	108	1	US-07-820-154A-32	Sequence 32, Appl
	24	25	36.2	108	3	US-08-097-554A-32	Sequence 32, Appl
	25	25	36.2	108	5	US-08-480-640A-32	Sequence 32, Appl
	26	25	36.2	108	6	PCT-US93-00324-32	Sequence 32, Appl
	27	25	36.2	111	3	US-08-097-554A-78	Sequence 78, Appl
	28	25	36.2	117	5	US-08-480-640A-78	Sequence 78, Appl
	29	25	36.2	178	4	US-08-484-575A-16	Sequence 16, Appl
	30	25	36.2	178	5	US-08-477-459-16	Sequence 16, Appl
	31	25	36.2	178	6	PCT-US94-01826A-16	Sequence 16, Appl
	32	25	36.2	178	6	PCT-US94-02252A-16	Sequence 16, Appl
	33	25	36.2	182	3	US-08-097-554A-64	Sequence 64, Appl
	34	25	36.2	182	4	US-08-484-575A-15	Sequence 15, Appl
	35	25	36.2	182	5	US-08-477-459-15	Sequence 15, Appl
	36	25	36.2	182	5	US-08-480-640A-64	Sequence 64, Appl
	37	25	36.2	182	6	PCT-US94-01826A-15	Sequence 15, Appl
	38	25	36.2	182	6	PCT-US94-02252A-15	Sequence 15, Appl
	39	25	36.2	185	5	US-08-480-640A-156	Sequence 156, App
	40	25	36.2	206	1	US-07-820-154A-21	Sequence 21, Appl
	41	25	36.2	206	3	US-08-097-554A-21	Sequence 21, Appl
	42	25	36.2	206	5	US-08-480-640A-21	Sequence 21, Appl
	43	25	36.2	206	6	PCT-US93-00324-21	Sequence 21, Appl
	44	25	36.2	4177	4	US-08-484-575A-12	Sequence 12, Appl
	45	25	36.2	4177	5	US-08-477-459-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-07-803-633A-6/c
; Sequence 6, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-803-633A-6

Query Match 38.3%; Score 26.4; DB 1; Length 40;
Best Local Similarity 83.3%; Pred. No. 0.39;
Matches 30; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGG 40
          ||||||||||||||||     |||||||||||| |||
Db     39 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4


RESULT   2
US-08-525-742-16/c
; Sequence 16, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-16


  Query Match             38.3%;  Score 26.4;  DB 3;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 0.39;
  Matches   30;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGG 40
          ||||||||||||||||     |||||||||||| |||
Db     39 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4


RESULT   3
US-08-525-742-30/c
; Sequence 30, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-30


  Query Match             38.3%;  Score 26.4;  DB 3;  Length 96;
  Best Local Similarity   83.3%;  Pred. No. 0.49;
  Matches   30;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGG 40
          ||||||||||||||||     |||||||||||| |||
Db     39 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4
```

```
RESULT   4
US-08-708-856A-3
; Sequence 3, Application US/08708856A
; Patent No. 5963254
;  GENERAL INFORMATION:
;    APPLICANT:  Ozaki, Akio
;    APPLICANT:  Mori, Hideo
;    APPLICANT:  Shibasaki, Takeshi
;    APPLICANT:  Ando, Katsuhiko
;    APPLICANT:  Ochiai, Keiko
;    APPLICANT:  Chiba, Shigeru
;    APPLICANT:  Uosaki, Youichi
;    TITLE OF INVENTION:  Process for Producing
;    TITLE OF INVENTION:  Cis-3-Hydroxy-L-Proline
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANTONELLI, TERRY, STOUT AND KRAUS
;      STREET:  1300 NORTH SEVENTEENTH STREET
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22209
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/708,856A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/301,654
;      FILING DATE:  07-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/474,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terry, David T.
;      REGISTRATION NUMBER:  20178
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-312-6600
;      TELEFAX:  703-312-6666
;      TELEX:  248545
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  870 base paires
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Streptomyces sp.
;      STRAIN:  TH1
US-08-708-856A-3


  Query Match            37.7%;  Score 26;  DB 4;  Length 870;
  Best Local Similarity   70.0%;  Pred. No. 1.2;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
           | ||||||  ||  |  |||||||||||| |    | |||||| ||||  ||
Db     441 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 490


RESULT   5
US-09-287-375-3
; Sequence 3, Application US/09287375
; Patent No. 6010891
;  GENERAL INFORMATION:
;    APPLICANT:  Ozaki, Akio
;    APPLICANT:  Mori, Hideo
;    APPLICANT:  Shibasaki, Takeshi
;    APPLICANT:  Ando, Katsuhiko
;    APPLICANT:  Ochiai, Keiko
;    APPLICANT:  Chiba, Shigeru
;    APPLICANT:  Uosaki, Youichi
;    TITLE OF INVENTION:  Process for Producing
;    TITLE OF INVENTION:  Cis-3-Hydroxy-L-Proline
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANTONELLI, TERRY, STOUT AND KRAUS
;      STREET:  1300 NORTH SEVENTEENTH STREET
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22209
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/287,375
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/708,856
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/474,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terry, David T.
;      REGISTRATION NUMBER:  20178
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-312-6600
;      TELEFAX:  703-312-6666
;      TELEX:  248545
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  870 base paires
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Streptomyces sp.
;      STRAIN:  TH1
US-09-287-375-3


  Query Match            37.7%;  Score 26;  DB 5;  Length 870;
  Best Local Similarity   70.0%;  Pred. No. 1.2;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
           | ||||||  ||  |  |||||||||||| |    | |||||| ||||  ||
Db     441 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 490


RESULT   6
US-08-708-856A-13
; Sequence 13, Application US/08708856A
; Patent No. 5963254
;  GENERAL INFORMATION:
;    APPLICANT:  Ozaki, Akio
;    APPLICANT:  Mori, Hideo
;    APPLICANT:  Shibasaki, Takeshi
```

```
;    APPLICANT:  Ando, Katsuhiko
;    APPLICANT:  Ochiai, Keiko
;    APPLICANT:  Chiba, Shigeru
;    APPLICANT:  Uosaki, Youichi
;    TITLE OF INVENTION:  Process for Producing
;    TITLE OF INVENTION:  Cis-3-Hydroxy-L-Proline
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANTONELLI, TERRY, STOUT AND KRAUS
;      STREET:  1300 NORTH SEVENTEENTH STREET
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22209
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/708,856A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/301,654
;      FILING DATE:  07-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/474,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terry, David T.
;      REGISTRATION NUMBER:  20178
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-312-6600
;      TELEFAX:  703-312-6666
;      TELEX:  248545
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base paires
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Streptomyces sp.
;      STRAIN:  TH1
US-08-708-856A-13


  Query Match             37.7%;  Score 26;  DB 4;  Length 1081;
  Best Local Similarity   70.0%;  Pred. No. 1.3;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
           | ||||||  ||  |  |||||||||||  |   | ||||||  ||||  ||
Db     491 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 540


RESULT   7
US-09-287-375-13
; Sequence 13, Application US/09287375
; Patent No. 6010891
;  GENERAL INFORMATION:
;    APPLICANT:  Ozaki, Akio
;    APPLICANT:  Mori, Hideo
;    APPLICANT:  Shibasaki, Takeshi
;    APPLICANT:  Ando, Katsuhiko
;    APPLICANT:  Ochiai, Keiko
;    APPLICANT:  Chiba, Shigeru
;    APPLICANT:  Uosaki, Youichi
;    TITLE OF INVENTION:  Process for Producing
;    TITLE OF INVENTION:  Cis-3-Hydroxy-L-Proline
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ANTONELLI, TERRY, STOUT AND KRAUS
;      STREET:  1300 NORTH SEVENTEENTH STREET
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22209
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/287,375
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/708,856
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/474,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terry, David T.
;      REGISTRATION NUMBER:  20178
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-312-6600
;      TELEFAX:  703-312-6666
;      TELEX:  248545
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base paires
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Streptomyces sp.
;      STRAIN:  TH1
US-09-287-375-13


  Query Match             37.7%;  Score 26;  DB 5;  Length 1081;
  Best Local Similarity   70.0%;  Pred. No. 1.3;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
           | ||||||  ||  |  |||||||||||  |   | ||||||  ||||  ||
Db     491 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 540


RESULT   8
US-08-525-742-28/c
; Sequence 28, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  93 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-28


  Query Match             37.4%;  Score 25.8;  DB 3;  Length 93;
  Best Local Similarity   81.1%;  Pred. No. 0.79;
  Matches   30;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGC 41
           ||||||||||||||||     |||||||||| || |
Db      37 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGTC 1


RESULT   9
US-08-525-742-27
; Sequence 27, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-27


  Query Match             37.4%;  Score 25.8;  DB 3;  Length 97;
  Best Local Similarity   81.1%;  Pred. No. 0.8;
  Matches   30;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGC 41
           ||||||||||||||||     |||||||||| || |
Db      61 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGTC 97


RESULT  10
US-07-803-633A-5
; Sequence 5, Application US/07803633A
; Patent No. 5369025
;  GENERAL INFORMATION:
;    APPLICANT:  NAZERIAN, Keyvan
;    APPLICANT:  LEE, Lucy F.
;    APPLICANT:  YANAGIDA, No. 5369025oru
;    APPLICANT:  OGAWA, Ryohei
;    APPLICANT:  LI, Yi
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VACCINE FOR
;    TITLE OF INVENTION:  PROTECTION AGAINST MAREK'S DISEASE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 No. 5369025th Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/803,633A
;      FILING DATE:  19911210
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1644-103P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 241-1300
;      TELEFAX:  (703) 241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-803-633A-5


  Query Match             36.8%;  Score 25.4;  DB 1;  Length 40;
  Best Local Similarity   82.9%;  Pred. No. 0.88;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
             |||||||||||||||     |||||||||||  ||
Db       6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


RESULT  11
US-08-525-742-15
; Sequence 15, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-15


  Query Match             36.8%;  Score 25.4;  DB 3;  Length 40;
  Best Local Similarity   82.9%;  Pred. No. 0.88;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
             |||||||||||||||     |||||||||||  ||
Db       6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


RESULT  12
US-08-525-742-29
; Sequence 29, Application US/08525742
; Patent No. 5871742
;  GENERAL INFORMATION:
;    APPLICANT:  Saito, Shuji
;    APPLICANT:  Ohkawa, Setsuko
;    APPLICANT:  Saeki, Sakiko
;    APPLICANT:  Ohsawa, Ikuroh
;    APPLICANT:  Funato, Hirono
;    APPLICANT:  Iritani, Yoshikazu
;    APPLICANT:  Aoyama, Shigemi
;    APPLICANT:  Takahashi, Kiyoohito
;    TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;    TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;    TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;    TITLE OF INVENTION:  AS USE THEREOF
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;      ADDRESSEE:  NAUGHTON
;      STREET:  1725 K Street, Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/525,742
;      FILING DATE:  25-SEP-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-074139
;      FILING DATE:  31-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-245625
;      FILING DATE:  30-SEP-1993
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  PCT/JP94/00541
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McLeland, Le-Nhung
;      REGISTRATION NUMBER:  31,541
;      REFERENCE/DOCKET NUMBER:  950811
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-659-2930
;      TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "synthetic DNA"
US-08-525-742-29


  Query Match             36.8%;  Score 25.4;  DB 3;  Length 95;
  Best Local Similarity   82.9%;  Pred. No. 1.1;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||||     |||||||||| ||
Db     61 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 95


RESULT  13
US-08-484-575A-8
; Sequence 8, Application US/08484575A
; Patent No. 5925358
;  GENERAL INFORMATION:
;    APPLICANT:  Mark D. Cochran and David E. Junker
;    TITLE OF INVENTION:  Recombinant Fowlpox Viruses and Uses Thereof
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,575A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White Esq, John P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)278-0450
;      TELEFAX:  (212)391-0525
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-484-575A-8


  Query Match             36.2%;  Score 25;  DB 4;  Length 40;
  Best Local Similarity   84.8%;  Pred. No. 1.2;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||     ||||||||||||
Db      1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


RESULT  14
US-08-477-459-8
; Sequence 8, Application US/08477459
; Patent No. 6001369
;  GENERAL INFORMATION:
;    APPLICANT:  Mark D. Cochran
;    TITLE OF INVENTION:  Recombinant Fowlpox Viruses and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,459
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White Esq, John P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-459-8


  Query Match             36.2%;  Score 25;  DB 5;  Length 40;
  Best Local Similarity   84.8%;  Pred. No. 1.2;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||     ||||||||||||
Db      1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


RESULT  15
PCT-US94-01826A-8
; Sequence 8, Application PC/TUS9401826A
;  GENERAL INFORMATION:
;    APPLICANT:  Syntro Corporation, et al.
;    TITLE OF INVENTION:  Recombinant Fowlpox Virus S-FPV-043 and Uses  Thereof
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/01826A
;       FILING DATE:  28-FEB-1994
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White Esq, John P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)977-9550
;       TELEFAX:  (212)664-0525
;       TELEX:  422523
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  40 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
PCT-US94-01826A-8


  Query Match             36.2%;  Score 25;  DB 6;  Length 40;
  Best Local Similarity   84.8%;  Pred. No. 1.2;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
           ||||||||||||||||     ||||||||||||
Db       1 AAAAATTGAAAAACTATTCTAATTTATTGCACG 33


Search completed: May 29, 2000, 22:08:49
Job time: 38730 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:21:08 ; Search time 1214.87 Seconds
(without alignments)
432.625 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GGCCAAAAATTGAAAAACTA.....GCCCGGCCGCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		28.6	41.4	2360	5	E08871	E08871 A part of g
2		28.6	41.4	4987	5	E08872	E08872 DNA encodin
3		27.8	40.3	207	5	E04525	E04525 Vaccinia vi
4		26.6	38.6	38	5	E07991	E07991 ATI promote
c 5		26.6	38.5	38	5	E07992	E07992 ATI promote
c 6		26.4	38.3	40	5	AR035285	AR035285 Sequence
c 7		26.4	38.3	96	5	AR035298	AR035298 Sequence
c 8		26.2	38.0	167879	32	AP000827	AP000827 Homo sapi
9		26	37.7	2900	1	AB007189	AB007189 Streptomy
c 10		26	37.7	186326	52	AC016688	AC016688 Homo sapi
c 11		25.8	37.4	93	5	AR035296	AR035296 Sequence
12		25.8	37.4	97	5	AR035295	AR035295 Sequence
13		25.4	36.8	36	24	E12061	E12061 Synthetic e
14		25.4	36.8	40	5	AR035284	AR035284 Sequence
15		25.4	36.8	40	5	E08848	E08848 Poxvirus ea
16		25.4	36.8	95	5	AR035297	AR035297 Sequence
17		25.2	36.5	3612	16	EDSORFS	Z86066 Eggdrop syn
18		25.2	36.5	33213	16	AAVEDSDNA	Y09598 Avian adeno
19		25	36.2	70	5	AR034185	AR034185 Sequence
20		25	36.2	102	5	AR034169	AR034169 Sequence
21		25	36.2	108	5	AR034179	AR034179 Sequence
22		25	36.2	111	5	AR034209	AR034209 Sequence
23		25	36.2	182	5	AR034200	AR034200 Sequence
24		25	36.2	206	5	AR034174	AR034174 Sequence
c 25		25	36.2	152453	42	AC012130	AC012130 Homo sapi
26		24.8	35.9	226	24	E12074	E12074 Synthetic p
27		24.8	35.9	1433	1	MTU43178	U43178 Mycobacteri
28		24.8	35.9	1580	16	AD5004	X02998 Adenovirus
29		24.8	35.9	1580	16	ADRA3	J01969 adenovirus
30		24.8	35.9	3189	5	A52460	A52460 Sequence 4
31		24.8	35.9	24171	34	CEC31A11	Z83218 Caenorhabdi
32		24.8	35.9	35935	16	ADRCOMPGEN	M73260 Mastadenovi
33		24.8	35.9	35937	16	ADRCG	J01917 Adenovirus
c 34		24.8	35.9	69316	55	AC023645	AC023645 Homo sapi
35		24.8	35.9	157904	56	AC022361	AC022361 Homo sapi
36		24.8	35.9	217109	53	AC018806	AC018806 Homo sapi
c 37		24.6	35.7	24595	1	ATACH5	X00493 Agrobacteri
c 38		24.6	35.7	24595	5	E00404	E00404 Ti plasmid
c 39		24.6	35.7	24595	5	E00546	E00546 DNA fragmen
40		24.6	35.7	27608	14	AF184978	AF184978 Binary ve
41		24.6	35.7	72496	44	AC019064	AC019064 Homo sapi
42		24.6	35.7	172749	40	AC009731	AC009731 Homo sapi
43		24.4	35.4	86	5	AR035318	AR035318 Sequence
44		24.4	35.4	91	5	AR035319	AR035319 Sequence
45		24.4	35.4	2080	2	AFARBCAE	M17744 Ralstonia e

ALIGNMENTS

```
RESULT   1
E08871
LOCUS       E08871       2360 bp    RNA             PAT       29-SEP-1997
DEFINITION  A part of genomic sequence of Vaccinia virus.
ACCESSION   E08871
VERSION     E08871.1  GI:2176975
KEYWORDS    JP 1995069899-A/3.
SOURCE      Vaccinia virus.
  ORGANISM  Vaccinia virus
            Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
            Orthopoxvirus.
REFERENCE   1  (bases 1 to 2360)
  AUTHORS   Seki,M., Honda,Y. and Yamada,S. .
  TITLE     ANTIVIRAL AGENT
  JOURNAL   Patent: JP 1995069899-A 3 14-MAR-1995;
            MITSUBISHI CHEM CORP
COMMENT     OS   Unknown (Vaccinia virus)
            PN   JP 1995069899-A/3
            PD   14-MAR-1995
            PF   02-SEP-1993 JP 1993241973
            PI   SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
            PC   A61K31/70,A61K48/00,C07K7/00,C12N15/09//C07H21/04; CC
            strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2360
            FT                   /organism='Unclassified'
            FT                   /clone='pHASE'.
FEATURES             Location/Qualifiers
     source          1. .2360
                     /organism="Vaccinia virus"
                     /db_xref="taxon:10245"
BASE COUNT      858 a    410 c    399 g    693 t
ORIGIN


  Query Match             41.4%;  Score 28.6;  DB 5;  Length 2360;
  Best Local Similarity   88.6%;  Pred. No. 12;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          ||||||||||||||||||    ||||||||||||
Db   1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269


RESULT   2
E08872
LOCUS       E08872       4987 bp    DNA             PAT       29-SEP-1997
DEFINITION  DNA encoding a fusion protein of Vaccinia virus protein,HCV protein
             and Firefly luciferase.
ACCESSION   E08872
VERSION     E08872.1  GI:2176976
KEYWORDS    JP 1995069899-A/4.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 4987)
  AUTHORS   Seki,M., Honda,Y. and Yamada,S. .
  TITLE     ANTIVIRAL AGENT
  JOURNAL   Patent: JP 1995069899-A 4 14-MAR-1995;
            MITSUBISHI CHEM CORP
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1995069899-A/4
            PD   14-MAR-1995
            PF   02-SEP-1993 JP 1993241973
            PI   SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
            PC   A61K31/70,A61K48/00,C07K7/00,C12N15/09//C07H21/04; CC
            strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .4987
            FT                   /organism='Artificial sequences' FT   CDS
              1826. .4057
            FT                   /product='fusion protein of virus protein and
            FT                   luciferase'.
FEATURES             Location/Qualifiers
     source          1. .4987
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT     1511 a   1041 c   1073 g   1314 t     48 others
ORIGIN


  Query Match             41.4%;  Score 28.6;  DB 5;  Length 4987;
  Best Local Similarity   88.6%;  Pred. No. 13;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          ||||||||||||||||||    ||||||||||||
Db   1235 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 1269


RESULT   3
E04525
LOCUS       E04525       207 bp    DNA             PAT       29-SEP-1997
DEFINITION  Vaccinia virus early promoter.
ACCESSION   E04525
VERSION     E04525.1  GI:2172726
KEYWORDS    JP 1993078395-A/6.
SOURCE      Vaccinia virus.
  ORGANISM  Vaccinia virus
            Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
            Orthopoxvirus.
REFERENCE   1  (bases 1 to 207)
  AUTHORS   Yasuda,K., Sato,T., Nagaya,A. and Kyo,T. .
  TITLE     GLUCOPROTEIN, ITS PRODUCTION, RECOMBINED VACCINIA VIRUS FOR ITS
            PRODUCTION, AND DIAGNOSTIC AGENT USING THE SAME
  JOURNAL   Patent: JP 1993078395-A 6 30-MAR-1993;
            NIPPON ZEON CO LTD
COMMENT     OS   Vaccinia virus
            PN   JP 1993078395-A/6
            PD   30-MAR-1993
            PF   20-SEP-1991 JP 1991241275
            PI   YASUDA KANJI, SATO TAKANORI, NAGAYA ATSUSHI, KYO TSUGUO PC
            C07K9/00,A61K39/00,A61K39/395,C07K15/14,C12N7/01,C12P21/00, PC
            G01N33/569,
            PC   G01N33/576//C12N15/51,(C12P21/00,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   promoter        1. .207
            FT                   /note='vaccinia virus early promoter'.
FEATURES             Location/Qualifiers
     source          1. .207
                     /organism="Vaccinia virus"
                     /db_xref="taxon:10245"
BASE COUNT       66 a     48 c     38 g     55 t
ORIGIN


  Query Match             40.3%;  Score 27.8;  DB 5;  Length 207;
  Best Local Similarity   74.5%;  Pred. No. 16;
  Matches   35;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGC 51
          |||||||||||||||||     ||||||||||   | |  |||||| |
Db     79 AAAAATTGAAAAACTATTCTAATTTATTGCAATAGTCGACCATGGCC 125


RESULT   4
```

```
E07991
LOCUS       E07991         38 bp    DNA             PAT       29-SEP-1997
DEFINITION  ATI promoter derived from Poxvirus.
ACCESSION   E07991
VERSION     E07991.1  GI:2176122
KEYWORDS    JP 1994237773-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Funahashi,S., Kin,Y. and Shida,H. .
  TITLE     A TYPE INCLUSION @(3754/24)ATI) PROMOTER OF POXVIRUS AND EXOGENOTE
            EXPRESSION VECTOR COMPRISING PROPHASE PROMOTER
  JOURNAL   Patent: JP 1994237773-A 1 30-AUG-1994;
            TONEN CORP, SHIDA HISATOSHI
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1994237773-A/1
            PD   30-AUG-1994
            PF   21-JAN-1993 JP 1993008424
            PR   22-DEC-1992 JP 92P   342362
            PI   FUNAHASHI SHINICHI, KIN YASUKAZU, SHIDA HISATOSHI PC
            C12N15/39,C12N7/01,(C12N15/39,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .38
            FT                   /organism='Artificial sequences'.
FEATURES              Location/Qualifiers
     source           1. .38
                      /organism="unidentified"
                      /db_xref="taxon:32644"
BASE COUNT       15 a      4 c      7 g     12 t
ORIGIN


  Query Match            38.6%;  Score 26.6;  DB 5;  Length 38;
  Best Local Similarity  87.9%;  Pred. No. 32;
  Matches   29;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||    ||||||||||||
Db      1 AAAAATTGAAAAACTAGTCTAATTTATTGCACG 33


RESULT   5
E07992/c
LOCUS       E07992         38 bp    DNA             PAT       29-SEP-1997
DEFINITION  ATI promoter derived from Poxvirus.
ACCESSION   E07992
VERSION     E07992.1  GI:2176123
KEYWORDS    JP 1994237773-A/2.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Funahashi,S., Kin,Y. and Shida,H. .
  TITLE     A TYPE INCLUSION @(3754/24)ATI) PROMOTER OF POXVIRUS AND EXOGENOTE
            EXPRESSION VECTOR COMPRISING PROPHASE PROMOTER
  JOURNAL   Patent: JP 1994237773-A 2 30-AUG-1994;
            TONEN CORP, SHIDA HISATOSHI
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1994237773-A/2
            PD   30-AUG-1994
            PF   21-JAN-1993 JP 1993008424
            PR   22-DEC-1992 JP 92P   342362
            PI   FUNAHASHI SHINICHI, KIN YASUKAZU, SHIDA HISATOSHI PC
            C12N15/39,C12N7/01,(C12N15/39,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .38
            FT                   /organism='Artificial sequences'.
FEATURES              Location/Qualifiers
     source           1. .38
                      /organism="unidentified"
                      /db_xref="taxon:32644"
BASE COUNT       12 a      7 c      4 g     15 t
ORIGIN


  Query Match            38.6%;  Score 26.6;  DB 5;  Length 38;
  Best Local Similarity  87.9%;  Pred. No. 32;
  Matches   29;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||    ||||||||||||
Db     34 AAAAATTGAAAAACTAGTCTAATTTATTGCACG 2


RESULT   6
AR035285/c
LOCUS       AR035285       40 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 16 from patent US 5871742.
ACCESSION   AR035285
VERSION     AR035285.1  GI:5951953
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 16 16-FEB-1999;
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="unknown"
BASE COUNT       14 a      4 c      6 g     16 t
ORIGIN


  Query Match            38.3%;  Score 26.4;  DB 5;  Length 40;
  Best Local Similarity  83.3%;  Pred. No. 38;
  Matches   30;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGG 40
          |||||||||'||||||    |||||||||||  |||
Db     39 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4


RESULT   7
AR035298/c
LOCUS       AR035298       96 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 30 from patent US 5871742.
ACCESSION   AR035298
VERSION     AR035298.1  GI:5951966
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 30 16-FEB-1999;
FEATURES              Location/Qualifiers
     source           1. .96
                      /organism="unknown"
BASE COUNT       44 a      8 c     10 g     34 t
ORIGIN
```

```
  Query Match             38.3%;  Score 26.4;  DB 5;  Length 96;
  Best Local Similarity   83.3%;  Pred. No. 43;
  Matches   30;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGG 40
           ||||||||||||||||     |||||||||||  |||
Db      39 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4


RESULT   8
AP000827/c
LOCUS       AP000827   167879 bp    DNA             HTG       04-FEB-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-716D19 map 11q14, WORKING
            DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION   AP000827
VERSION     AP000827.1  GI:6997662
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens DNA, clone:RP11-716D19.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167879)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 167,879 genomic DNA of 11q14
  JOURNAL   Published Only in DataBase (1999) In press
REFERENCE   2  (bases 1 to 167879)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-1999) to the DDBJ/EMBL/GenBank databases.
            Masahira Hattori, The Institute of Physical and Chemical Research
            (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1
            Kitasato, Sagamihara, Kanagawa 228-8555, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-42-778-9923, Fax:81-42-778-9924)
COMMENT     On Feb 18, 2000 this sequence version replaced gi:6525269.
            -------------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
                Center project name: HumDraft11
                Center clone name: RP11-716D19
            -------------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 140881 bases at least Q40
                Consensus quality: 149365 bases at least Q30
                Consensus quality: 154204 bases at least Q20
                Insert size: 156754; sum-of-contigs
                Quality coverage: 4.32x in Q20 bases; sum-of-contigs
            --------------.
            NOTE: This is a 'working draft' sequence. It currently consists of
            24 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
                  1     14626 contig of  14626 of in length
              15127     30176 contig of  15050 of in length
              30677     45326 contig of  14650 of in length
              45827     58556 contig of  12730 of in length
              59057     66715 contig of   7659 of in length
              67216     77376 contig of  10161 of in length
              77877     88355 contig of  10479 of in length
              88856     96474 contig of   7619 of in length
              96975    105697 contig of   8723 of in length
             106198    114759 contig of   8562 of in length
             115260    122523 contig of   7264 of in length
             123024    130348 contig of   7325 of in length
             130849    137848 contig of   7000 of in length
             138349    142439 contig of   4091 of in length
             142940    146239 contig of   3300 of in length
             146740    150037 contig of   3298 of in length
             150538    153200 contig of   2663 of in length
             153701    155617 contig of   1917 of in length
             156118    157980 contig of   1863 of in length
             158481    160479 contig of   1999 of in length
             160980    162394 contig of   1415 of in length
             162895    164348 contig of   1454 of in length
             164849    166183 contig of   1335 of in length
             166684    167879 contig of   1196 of in length.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 24 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1    14626: contig of 14626 bp in length
          *   14627 15130: gap of      504 bp
          *   15131    30176: contig of 15046 bp in length
          *   30177 30677: gap of      501 bp
          *   30678    45326: contig of 14649 bp in length
          *   45327 45831: gap of      505 bp
          *   45832    58551: contig of 12720 bp in length
          *   58552 59056: gap of      505 bp
          *   59057    66714: contig of 7658 bp in length
          *   66715 67215: gap of      501 bp
          *   67216    77375: contig of 10160 bp in length
          *   77376 77876: gap of      501 bp
          *   77877    88354: contig of 10478 bp in length
          *   88355 88855: gap of      501 bp
          *   88856    96467: contig of 7612 bp in length
          *   96468 96975: gap of      508 bp
          *   96976   105693: contig of 8718 bp in length
          *  105694 106198: gap of      505 bp
          *  106199   114752: contig of 8554 bp in length
          *  114753 115262: gap of      510 bp
          *  115263   122523: contig of 7261 bp in length
          *  122524 123027: gap of      504 bp
          *  123028   130348: contig of 7321 bp in length
          *  130349 130849: gap of      501 bp
          *  130850   137846: contig of 6997 bp in length
          *  137847 138349: gap of      503 bp
          *  138350   142439: contig of 4090 bp in length
          *  142440 142939: gap of      500 bp
          *  142940   146235: contig of 3296 bp in length
          *  146236 146739: gap of      504 bp
          *  146740   150031: contig of 3292 bp in length
          *  150032 150537: gap of      506 bp
          *  150538   153199: contig of 2662 bp in length
          *  153200 153702: gap of      503 bp
          *  153703   155615: contig of 1913 bp in length
          *  155616 156117: gap of      502 bp
          *  156118   157978: contig of 1861 bp in length
          *  157979 158480: gap of      502 bp
          *  158481   160478: contig of 1998 bp in length
          *  160479 160981: gap of      503 bp
          *  160982   162383: contig of 1402 bp in length
          *  162384 162894: gap of      511 bp
          *  162895   164341: contig of 1447 bp in length
          *  164342 164848: gap of      507 bp
          *  164849   166183: contig of 1335 bp in length
          *  166184 166685: gap of      502 bp
          *  166686   167879: contig of 1194 bp in length.
FEATURES             Location/Qualifiers
     source          1. .167879
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP11-716D19"
                     /map="11q14"
BASE COUNT    49210 a  28817 c  28130 g  48047 t  13675 others
ORIGIN


  Query Match             38.0%;  Score 26.2;  DB 32;  Length 167879;
  Best Local Similarity   90.3%;  Pred. No. 1.5e+02;
  Matches   28;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTAT 31
          ||| |||||||| |||||||||||||||| |
Db  74276 GGCTAAAAATTGCAAAACTAGATCTATTTGT 74246


RESULT   9
AB007189
LOCUS       AB007189     2900 bp    DNA             BCT       27-SEP-1997
DEFINITION  Streptomyces sp. DNA for L-proline 3-hydroxylase, complete cds.
ACCESSION   AB007189 D78337
VERSION     AB007189.1  GI:2443305
KEYWORDS    L-proline 3-hydroxylase.
SOURCE      Streptomyces sp. (isolate:TH1) DNA, clone:pTH30.
  ORGANISM  Streptomyces sp.
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 2900)
  AUTHORS   Shibasaki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-1997) to the DDBJ/EMBL/GenBank databases. Takeshi
            Shibasaki, Kyowa Hakko Kogyo Co., Ltd., Tokyo Research Laboratory;
            3-6-6 Asahimachi, Machida, Tokyo 194, Japan (Tel:0427-25-2555,
            Fax:0427-26-8330)
REFERENCE   2  (sites)
  AUTHORS   Mori,H., Shibasaki,T., Yano,K. and Ozaki,A.
  TITLE     Purification and cloning of a proline 3-hydroxylase, a novel enzyme
            which hydroxylates free L-proline to cis-3-hydroxy-L-proline
  JOURNAL   J. Bacteriol. 179 (18), 5677-5683 (1997)
  MEDLINE   97440113
COMMENT     On Sep 27, 1997 this sequence version replaced gi:1747311.
            D78337: submitted (17-Nov-1995).
FEATURES             Location/Qualifiers
     source          1. .2900
                     /organism="Streptomyces sp."
                     /isolate="TH1"
                     /db_xref="taxon:1931"
                     /clone="pTH30"
     CDS             297. .1208
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA22405.1"
                     /db_xref="GI:2443306"
                     /translation="MQPLGDHAPDPDGDLPAELGVLCAQLPHGSGVELPXPHRAHRLG
                     AQVTEVRREVPGPADQFTVTERLDDGGPAAGHVQVEGDGTGAQDPEAVGAPPLLEQAG
                     PGLVGDLTGEAHQPFAVHGLKSRGHGGDGGHRTGSLHGFCSQRGRTPLGAPRATRARP
                     GGFRPHQVGRVALRSTSGALRIPSPPPPRDSFRLNNPAARQETVDCQDAAKLGSIFEC
                     GSELGSCSAATEGDWGDLPFAECRRPEHAHRVCLGFFVMSDHRPFSGRLLRVHLTFGY
                     RLLALALPVAILDASHRECGGDSRLPR"
     CDS             1312. .2184
                     /codon_start=1
                     /transl_table=11
                     /product="L-proline 3-hydroxylase"
                     /protein_id="BAA22406.1"
                     /db_xref="GI:2443307"
                     /translation="MRSHILGRIELDQERLGRDLEYLATVPTVEEEYDEFSNGFWKNI
                     PLYNASGGSEDRLYRDLEGSPAQPTKHAEQVPYLNEIITTVYNGERLQMARTRNLKNA
                     VVIPHRDFVELDRELDQYFRTHLMLEDSPLAFHSDDDTVIHMRAGEIWFLDAAAVHSA
                     VNFAEFSRQSLCVDLAFDGAFDEKEAFADATVYAPNLSPDVRERKPFTKEREAGILAL
                     SGVIGRENFRDILFLLSKVHYTYDVHPGETFEWLVSVSKGAGDDKMVEKAERIRDFAI
                     GARALGERFSLTTW"

     CDS             2244. .2870
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA22407.1"
                     /db_xref="GI:2443308"
                     /translation="MSAEAFDDDADIQLFAEERAAVADAVPRRRREFSTVRRCARAAL
                     GELGIPPVPLPPGRHRAPQWPTGVVGSMTHCSGYRAAAVARASRLHSVGIDAEESAPL
                     PDGVLDLVGLPXERDQVERLGAQSDAVPWDRLLFSCKEAVYKVWFPLAQRLLGFDGAR
                     IDIDSGGFFSARFLVPPPLASGNPVPRLTGRWMHRDGLILTAIALPCP"
BASE COUNT      486 a    999 c    945 g    464 t      6 others
ORIGIN


  Query Match             37.7%;  Score 26;  DB 1;  Length 2900;
  Best Local Similarity   70.0%;  Pred. No. 96;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     18 CTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCG 67
          | ||||||  ||  |  |||||||||||| |   | |||||| ||||  ||
Db   1752 CGAGATCTGGTTCCTCGACGCGGCCGCCGTCCACTCGGCCGTCAACTTCG 1801


RESULT  10
AC016688/c
LOCUS       AC016688   186326 bp    DNA             HTG       31-JAN-2000
DEFINITION  Homo sapiens clone RP11-80J14, WORKING DRAFT SEQUENCE, 20 unordered
            pieces.
ACCESSION   AC016688
VERSION     AC016688.2  GI:6838910
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 186326)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186326)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jan 31, 2000 this sequence version replaced gi:6524391.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_NH0080J14
            ---------- Summary Statistics ----------
            Sequencing vector: M13; 80%
            Sequencing vector: plasmid; 20%
            Chemistry: Dye-primer ET; 80% of reads
            Chemistry: Dye-terminator Big Dye; 20% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 174077 bases at least Q40
            Consensus quality: 178110 bases at least Q30
            Consensus quality: 180386 bases at least Q20
            Insert size: 174000; agarose-fp
            Insert size: 186326; sum-of-contigs
            Quality coverage: 4.31 in Q20 bases; agarose-fp
            Quality coverage: 4.03 in Q20 bases; sum-of-contigs
            ----------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
```

```
            * as soon as it is available and the accession number will
            * be preserved.
            *         1     1456: contig of 1456 bp in length
            *                     gap of unknown length
            *      1457     2971: contig of 1515 bp in length
            *                     gap of unknown length
            *      2972     5544: contig of 2573 bp in length
            *                     gap of unknown length
            *      5545     7248: contig of 1704 bp in length
            *                     gap of unknown length
            *      7249     8398: contig of 1150 bp in length
            *                     gap of unknown length
            *      8399    10354: contig of 1956 bp in length
            *                     gap of unknown length
            *     10355    12998: contig of 2644 bp in length
            *                     gap of unknown length
            *     12999    16483: contig of 3485 bp in length
            *                     gap of unknown length
            *     16484    20173: contig of 3690 bp in length
            *                     gap of unknown length
            *     20174    27635: contig of 7462 bp in length
            *                     gap of unknown length
            *     27636    37720: contig of 10085 bp in length
            *                     gap of unknown length
            *     37721    46381: contig of 8661 bp in length
            *                     gap of unknown length
            *     46382    55849: contig of 9468 bp in length
            *                     gap of unknown length
            *     55850    63824: contig of 7975 bp in length
            *                     gap of unknown length
            *     63825    76918: contig of 13094 bp in length
            *                     gap of unknown length
            *     76919    90722: contig of 13804 bp in length
            *                     gap of unknown length
            *     90723   109347: contig of 18625 bp in length
            *                     gap of unknown length
            *    109348   126404: contig of 17057 bp in length
            *                     gap of unknown length
            *    126405   148648: contig of 22244 bp in length
            *                     gap of unknown length
            *    148649   186326: contig of 37678 bp in length.
FEATURES             Location/Qualifiers
     source          1. .186326
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-80J14"
BASE COUNT    54067 a  40241 c  39737 g  52239 t     42 others
ORIGIN


  Query Match             37.7%;  Score 26;  DB 52;  Length 186326;
  Best Local Similarity   62.1%;  Pred. No. 1.7e+02;
  Matches   41;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAA 62
           | ||| | | |||  ||   | | ||||  ||||  |||||||    ||||  |  ||||
Db    5570 CAAAAGAATAAAACTCTTCTTTTTTTTTAAAGCACAGGGCCGCCCCATGCCCCTCTCCCAA 5511

Qy      63 CGGCGG 68
           ||| |
Db    5510 GGGCCG 5505


RESULT  11
AR035296/c
LOCUS       AR035296       93 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 28 from patent US 5871742.
ACCESSION   AR035296
VERSION     AR035296.1  GI:5951964
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 28 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="unknown"
BASE COUNT       43 a      7 c     10 g     33 t
ORIGIN


  Query Match             37.4%;  Score 25.8;  DB 5;  Length 93;
  Best Local Similarity   81.1%;  Pred. No. 69;
  Matches   30;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGC 41
           |||||||||||||||     |||||||||||  || |
Db      37 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGTC 1


RESULT  12
AR035295
LOCUS       AR035295       97 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 27 from patent US 5871742.
ACCESSION   AR035295
VERSION     AR035295.1  GI:5951963
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 27 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="unknown"
BASE COUNT       34 a     11 c      8 g     44 t
ORIGIN


  Query Match             37.4%;  Score 25.8;  DB 5;  Length 97;
  Best Local Similarity   81.1%;  Pred. No. 69;
  Matches   30;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGC 41
           |||||||||||||||     |||||||||||  || |
Db      61 AAAAATTGAAAAACTATTCTAATTTATTGCACTCGTC 97


RESULT  13
E12061
ID   E12061     standard; DNA; UNC; 36 BP.
XX
AC   E12061;
XX
SV   E12061.1
XX
DT   08-OCT-1997 (Rel. 52, Created)
DT   08-OCT-1997 (Rel. 52, Last updated, Version 1)
XX
DE   Synthetic early promoter sequence.
XX
KW   JP 1996242869-A/2.
XX
OS   unidentified
OC   unclassified.
XX
RN   [1]
```

```
RP   1-36
RA   Kamogawa K., Ogawa R., Yamaguchi T., Hirai K.;
RT   "RECOMBINANT POXVIRUS AND VACCINE COMPRISING THE SAME";
RL   Patent number JP 1996242869-A/2, 24-SEP-1996.
RL   NIPPON ZEON CO LTD.
XX
CC   OS   None
CC   OC   Artificial sequences.
CC   PN   JP 1996242869-A/2
CC   PD   24-SEP-1996
CC   PF   16-JAN-1996 JP 1996023125
CC   PR   13-JAN-1995 JP 95P    21249
CC   PI   KAMOGAWA KOICHI, OGAWA RYOHEI, YAMAGUCHI TAKESHI,
CC   PI   HIRAI KATSUYA
CC   PC   C12N15/09,A61K39/12,C12N7/00;
CC   CC   strandedness: Double;
CC   CC   topology: Linear;
CC   CC   hypothetical: No;
CC   CC   anti-sense: No;
CC   FH   Key             Location/Qualifiers
CC   FH
CC   FT   source          1. .36
CC   FT                   /organism="Artificial sequences"
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .36
FT                   /db_xref="taxon:32644"
FT                   /organism="unidentified"
XX
SQ   Sequence 36 BP; 15 A; 5 C; 3 G; 13 T; 0 other;


  Query Match             36.8%;  Score 25.4;  DB 24;  Length 36;
  Best Local Similarity   82.9%;  Pred. No. 82;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||     |||||||||| ||
Db      2 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 36


RESULT  14
AR035284
LOCUS       AR035284       40 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 15 from patent US 5871742.
ACCESSION   AR035284
VERSION     AR035284.1  GI:5951952
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE     Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 15 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT       15 a      7 c      5 g     13 t
ORIGIN


  Query Match             36.8%;  Score 25.4;  DB 5;  Length 40;
  Best Local Similarity   82.9%;  Pred. No. 84;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||     |||||||||| ||
Db      6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


RESULT  15
E08848
LOCUS       E08848         40 bp    DNA             PAT       29-SEP-1997
DEFINITION  Poxvirus early promoter.
ACCESSION   E08848
VERSION     E08848.1  GI:2176952
KEYWORDS    JP 1995067655-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Yamaguchi,T., Fukushi,H., Hirai,K., Aoyama,S., Yamaguchi,T.,
            Iritani,K., Hayashi,Y., Ogawa,R., Takamura,C. and Kamogawa,K.
  TITLE     PROMOTER FOR COMBINED POXVIRUS AND RECOMBINENT POXVIRUS HAVING THE
  JOURNAL   Patent: JP 1995067655-A 1 14-MAR-1995;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1995067655-A/1
            PD   14-MAR-1995
            PF   30-AUG-1993 JP 1993238953
            PI   YAMAGUCHI TAKESHI, FUKUSHI HIDETO, HIRAI KATSUYA, PI   AOYAMA
            SHIGEMI,
            PI   YAMAGUCHI TAKESHI, IRITANI KOICHI, HAYASHI YUKIHIRO, PI
            OGAWA RYOHEI,
            PI   TAKAMURA CHIZUKO, KAMOGAWA KOICHI
            PC   C12N15/09,A61K39/275,C12N7/00,(C12N15/09,C12R1:92),(C12N7/00,
            PC   C12R1:92);
            CC   strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .40
            FT                   /organism='Artificial sequences'.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       15 a      7 c      5 g     13 t
ORIGIN


  Query Match             36.8%;  Score 25.4;  DB 5;  Length 40;
  Best Local Similarity   82.9%;  Pred. No. 84;
  Matches   29;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCG 39
          |||||||||||||||     |||||||||| ||
Db      6 AAAAATTGAAAAACTATTCTAATTTATTGCACTCG 40


Search completed: May 29, 2000, 21:33:53
Job time: 36765 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 17:05:30 ; Search time 2276.24 Seconds
(without alignments)
135.250 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GGCCAAAAATTGAAAAACTA.....GCCCGGCCGCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	69	100.0	69	28	US-08-935-377-1	Sequence 1, Appli
2	57	82.6	57	28	US-08-935-377-5	Sequence 5, Appli

	3	53	76.8	53	28	US-08-935-377-23 Sequence 23, Appl
c	4	53	76.8	53	28	US-08-935-377-24 Sequence 24, Appl
	5	49	71.0	145	28	US-08-935-377-6 Sequence 6, Appli
	6	49	71.0	148	28	US-08-935-377-7 Sequence 7, Appli
	7	48.8	70.7	149	28	US-08-935-377-8 Sequence 8, Appli
	8	48	69.6	150	28	US-08-935-377-9 Sequence 9, Appli
	9	37	53.6	75	28	US-08-935-377-3 Sequence 3, Appli
	10	36	52.2	70	28	US-08-935-377-30 Sequence 30, Appl
c	11	32	46.4	71	28	US-08-935-377-31 Sequence 31, Appl
	12	28.6	41.4	40	12	US-08-301-013-27 Sequence 27, Appl
	13	28.6	41.4	40	17	US-08-592-610-406 Sequence 406, App
	14	28.6	41.4	40	47	US-09-295-309A-406 Sequence 406, App
	15	28.6	41.4	234	13	US-08-358-928-89 Sequence 89, Appl
	16	28.6	41.4	234	18	US-08-651-472-89 Sequence 89, Appl
	17	28.6	41.4	2360	12	US-08-301-013-30 Sequence 30, Appl
	18	28.6	41.4	2360	17	US-08-592-610-409 Sequence 409, App
	19	28.6	41.4	2360	47	US-09-295-309A-409 Sequence 409, App
	20	28.6	41.4	4987	12	US-08-301-013-31 Sequence 31, Appl
	21	28.6	41.4	4987	17	US-08-592-610-410 Sequence 410, App
	22	28.6	41.4	4987	47	US-09-295-309A-410 Sequence 410, App
c	23	27.6	40.0	40	47	US-09-295-309A-411 Sequence 411, App
c	24	27.2	39.4	1341	57	US-09-514-000-7200 Sequence 7200, Ap
	25	27.2	39.4	47677	57	US-09-514-000-380 Sequence 380, App
c	26	26.6	38.6	14103	86	US-60-144-351-2441 Sequence 2441, Ap
c	27	26.4	38.3	40	3	US-07-722-860-6 Sequence 6, Appli
c	28	26.4	38.3	40	9	US-08-213-449A-6 Sequence 6, Appli
c	29	26.4	38.3	40	9	US-08-213-449B-6 Sequence 6, Appli
	30	26.4	38.3	101	37	US-09-055-293-3 Sequence 3, Appli
	31	26.2	38.0	678	40	US-09-107-532-2649 Sequence 2649, Ap
c	32	26.2	38.0	7206	69	US-60-050-444-687 Sequence 687, App
c	33	26.2	38.0	7206	70	US-60-068-186-691 Sequence 691, App
	34	26	37.7	870	16	US-08-474-135-3 Sequence 3, Appli
	35	26	37.7	870	19	US-08-708-856-3 Sequence 3, Appli
	36	26	37.7	870	55	US-09-455-406-3 Sequence 3, Appli
	37	26	37.7	1080	19	US-08-708-856-13 Sequence 13, Appl
	38	26	37.7	1081	16	US-08-474-135-11 Sequence 11, Appl
	39	26	37.7	1081	55	US-09-455-406-13 Sequence 13, Appl
	40	25.4	36.8	40	3	US-07-722-860-5 Sequence 5, Appli
	41	25.4	36.8	40	9	US-08-213-449A-5 Sequence 5, Appli
	42	25.4	36.8	40	9	US-08-213-449B-5 Sequence 5, Appli
	43	25.4	36.8	95	9	US-08-213-449B-20 Sequence 20, Appl
	44	25	36.2	40	16	US-08-479-869-8 Sequence 8, Appli
	45	25	36.2	40	16	US-08-484-575-8 Sequence 8, Appli

ALIGNMENTS

```
RESULT   1
US-08-935-377-1
; Sequence 1, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  46..69
US-08-935-377-1


  Query Match           100.0%;  Score 69;  DB 28;  Length 69;
  Best Local Similarity  100.0%;  Pred. No. 5.2e-15;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60

Qy      61 AACGGCGGA 69
           |||||||||
Db      61 AACGGCGGA 69


RESULT   2
US-08-935-377-5
; Sequence 5, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-5
```

```
  Query Match             82.6%;  Score 57;  DB 28;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-11;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCC 57
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCC 57


RESULT   3
US-08-935-377-23
; Sequence 23, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-23


  Query Match             76.8%;  Score 53;  DB 28;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCC 53
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCC 53


RESULT   4
US-08-935-377-24/c
; Sequence 24, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-24


  Query Match             76.8%;  Score 53;  DB 28;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCC 57
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     53 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCC 1


RESULT   5
US-08-935-377-6
; Sequence 6, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  cDNA
US-08-935-377-6


  Query Match             71.0%;  Score 49;  DB 28;  Length 145;
  Best Local Similarity   84.6%;  Pred. No. 9.5e-08;
  Matches   55;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
          ||||||||||||||||||||||||||||||||||||||||||| |||  ||  | | |
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGC 60

Qy     61 AACGG 65
             | |
Db     61 TGCAG 65


RESULT   6
US-08-935-377-7
; Sequence 7, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  148 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-7


  Query Match             71.0%;  Score 49;  DB 28;  Length 148;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-08;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGG 49
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGG 49


RESULT   7
US-08-935-377-8
; Sequence 8, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-8


  Query Match             70.7%;  Score 48.8;  DB 28;  Length 149;
  Best Local Similarity   88.3%;  Pred. No. 1.1e-07;
  Matches   53;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
          |||||||||||||||||||||||||||||||||||||||||||||| |     || ||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60


RESULT   8
US-08-935-377-9
; Sequence 9, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-9


  Query Match             69.6%;  Score 48;  DB 28;  Length 150;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-07;
  Matches   48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG 48
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG 48


RESULT  9
US-08-935-377-3
; Sequence 3, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  52..75
US-08-935-377-3


  Query Match             53.6%;  Score 37;  DB 28;  Length 75;
  Best Local Similarity   78.7%;  Pred. No. 0.0016;
  Matches   59;  Conservative    0;  Mismatches   10;  Indels    6;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTG------CACGCGGCCGCCATGGGCCCG 54
          ||||||||||||||||   ||  | | ||| |||       | ||||||||||||||||||
Db      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60

Qy     55 GCCGCCAACGGCGGA 69
          |||||||||||||||
Db     61 GCCGCCAACGGCGGA 75


RESULT  10
US-08-935-377-30
; Sequence 30, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  70 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-30


  Query Match             52.2%;  Score 36;  DB 28;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 0.0035;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCAC 36
          |||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCAC 36


RESULT  11
US-08-935-377-31/c
; Sequence 31, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  71 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-31


  Query Match             46.4%;  Score 32;  DB 28;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 0.095;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCAC 36
          ||||||||||||||||||||||||||||||||
Db     71 AAAAATTGAAAAACTAGATCTATTTATTGCAC 40


RESULT  12
US-08-301-013-27
; Sequence 27, Application US/08301013
;  GENERAL INFORMATION:
;    APPLICANT:  Makoto SEKI et al.
;    TITLE OF INVENTION:  ANTIVIRAL AGENT
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  805 Fifteenth Street, N.W., #700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 500 kb
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/301,013
;      FILING DATE:  September 6, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/299,113
;      FILING DATE:  September 2, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warren M. Cheek, Jr.
;      REGISTRATION NUMBER:  33,367
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-8850
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid synthetic DNA
US-08-301-013-27


  Query Match             41.4%;  Score 28.6;  DB 12;  Length 40;
  Best Local Similarity   88.6%;  Pred. No. 1.3;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||||    ||||||||||||
Db      4 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 38


RESULT  13
US-08-592-610-406
; Sequence 406, Application US/08592610
;  GENERAL INFORMATION:
;    APPLICANT:  Makoto SEKI et al.
;    TITLE OF INVENTION:  ANTISENSE COMPOUNDS COMPLEMENTARY TO HCV
;    TITLE OF INVENTION:  GENOME
;    NUMBER OF SEQUENCES:  410
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  805 Fifteenth Street, N.W., #700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 mb
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,610
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/111,465
;      FILING DATE:  August 25, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Matthew Jacob
;      REGISTRATION NUMBER:  25,154
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:
;  INFORMATION FOR SEQ ID NO:  406:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid, synthetic DNA
US-08-592-610-406


  Query Match             41.4%;  Score 28.6;  DB 17;  Length 40;
  Best Local Similarity   88.6%;  Pred. No. 1.3;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          |||||||||||||||||||    ||||||||||||
Db      4 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 38


RESULT  14
US-09-295-309A-406
; Sequence 406, Application US/09295309A
;  GENERAL INFORMATION:
;    APPLICANT:  Makoto SEKI et al.
;    TITLE OF INVENTION:  ANTISENSE COMPOUNDS COMPLEMENTARY TO HCV
;    TITLE OF INVENTION:  GENOME
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Wenderoth, Lind & Ponack, L.L.P.
;      STREET:  2033 K Street, N.W., Suite 800
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  U.S.A.
;      ZIP:  20006
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 mb
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect 5.1+
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/295,309A
;      FILING DATE:  April 21, 1999
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Matthew Jacob
;      REGISTRATION NUMBER:  25,154
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 721-8200
;      TELEFAX:  (202) 721-8250
;  INFORMATION FOR SEQ ID NO:  406:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid, synthetic DNA
US-09-295-309A-406


  Query Match             41.4%;  Score 28.6;  DB 47;  Length 40;
  Best Local Similarity   88.6%;  Pred. No. 1.3;
  Matches   31;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACG 37
          ||||||||||||||||||    ||||||||||||
Db      4 CCAAAAATTGAAAAACTAGTCTAATTTATTGCACG 38


RESULT  15
US-08-358-928-89
; Sequence 89, Application US/08358928
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  SCHEIFLINGER, Friedrich
;    APPLICANT:  FALKNER, Falko Gunter
;    APPLICANT:  PFLEIDERER, Michael
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;    TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;    TITLE OF INVENTION:  (HIV-1) ANTIGENS
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/358,928
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/914,738
;      FILING DATE:  20-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-358-928-89


  Query Match             41.4%;  Score 28.6;  DB 13;  Length 234;
  Best Local Similarity   72.5%;  Pred. No. 2.2;
  Matches   37;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGG 55
          |||||||||||||||||     ||||||||||||   |   |   |||||||
Db    183 AAAAATTGAAAAACTATTCTAATTTATTGCACGGTACGTACCATGGCCCGG 233


Search completed: May 30, 2000, 09:48:19
Job time: 60169 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:04:41 ; Search time 2192.43 Seconds
(without alignments)
127.563 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GGCCAAAAATTGAAAAACTA.....GCCCGGCCGCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	27	39.1	223	71	AV281309	AV281309 AV281309
c	2	26.6	38.6	464	36	AA606654	AA606654 vm93a06.r
	3	25.8	37.4	692	105	AQ274419	AQ274419 mgxb0015M
	4	25.6	37.1	1201	83	CNS0164L	AL106287 Drosophil
	5	25	36.2	492	106	AQ406794	AQ406794 HS_5103_B
	6	25	36.2	507	99	AQ212765	AQ212765 HS_3115_B
	7	25	36.2	516	85	AQ699107	AQ699107 HS_5560_B
	8	25	36.2	603	106	AQ415486	AQ415486 RPCI-11-1
	9	25	36.2	658	106	AQ384188	AQ384188 RPCI11-12
c	10	24.8	35.9	426	79	AW265066	AW265066 xq60g03.x
c	11	24.8	35.9	630	79	AW276836	AW276836 xp66g11.x
	12	24.6	35.7	379	48	C84505	C84505 C84505 oste
	13	24.4	35.4	194	80	AW354184	AW354184 32589 MAR
c	14	24.4	35.4	355	96	AQ064961	AQ064961 HS_2225_A
c	15	24.4	35.4	578	36	C78547	C78547 C78547 Mous
	16	24.2	35.1	221	32	AA339014	AA339014 EST44246
c	17	24.2	35.1	496	100	AQ146985	AQ146985 HS_2250_A
	18	24	34.8	270	70	AW143615	AW143615 EST293911
c	19	24	34.8	560	106	AQ398078	AQ398078 mgxb0010E
	20	23.8	34.5	329	21	T62402	T62402 tgEST0166 T
	21	23.8	34.5	336	27	AA012360	AA012360 TgESTzz17
	22	23.8	34.5	338	26	W96580	W96580 TgESTzy98h0
	23	23.8	34.5	351	27	AA011912	AA011912 TgESTzz10
	24	23.8	34.5	406	25	N81658	N81658 TgESTzy61h0
	25	23.8	34.5	414	27	AA012016	AA012016 TgESTzz17
	26	23.8	34.5	418	25	W06431	W06431 TgESTzy81a0
	27	23.8	34.5	420	25	W00132	W00132 TgESTzy74a0
	28	23.8	34.5	439	25	N60068	N60068 TgESTzy10e1
	29	23.8	34.5	441	25	N60176	N60176 TgESTzy03e1
	30	23.8	34.5	453	25	N59993	N59993 TgESTzy03h0
c	31	23.8	34.5	490	80	AW287387	AW287387 LG1_303_C
c	32	23.8	34.5	502	40	AA918332	AA918332 on42e03.s
	33	23.8	34.5	573	34	AA519942	AA519942 TgESTzz58
	34	23.8	34.5	579	34	AA520329	AA520329 TgESTzz69
c	35	23.8	34.5	584	85	AQ722468	AQ722468 HS_5228_A
	36	23.6	34.2	140	40	AA978155	AA978155 ny36b06.s
c	37	23.6	34.2	276	72	AV324511	AV324511 AV324511
	38	23.6	34.2	303	32	AA361896	AA361896 EST71242
	39	23.6	34.2	320	32	AA337828	AA337828 EST42610
	40	23.6	34.2	346	47	AI544871	AI544871 fb69c03.x
c	41	23.6	34.2	400	37	AA729295	AA729295 nx59c10.s
c	42	23.6	34.2	439	33	AA443577	AA443577 zw35a01.s
	43	23.6	34.2	442	54	AI721472	AI721472 fc28b03.x
	44	23.6	34.2	467	27	AA056924	AA056924 zk79c08.r
	45	23.6	34.2	506	90	AQ808912	AQ808912 HS_5252_A

ALIGNMENTS

RESULT 1
AV281309/c
LOCUS AV281309 223 bp mRNA EST 05-NOV-1999
DEFINITION AV281309 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cDNA clone 4933423P14 3', mRNA sequence.
ACCESSION AV281309
VERSION AV281309.1 GI:6269346
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 223)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5434327.
Contact: Yoshihide Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center, Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rtc.riken.go.jp,
URL:http://genome.rtc.riken.go.jp/
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .223
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4933423P14"
/clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
BASE COUNT 65 a 51 c 32 g 75 t
ORIGIN

Query Match 39.1%; Score 27; DB 71; Length 223;
Best Local Similarity 70.6%; Pred. No. 60;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGC 51
          |||||||||| |||||||| | | || ||  |  | | ||  || | ||||
Db     68 GGCCAAAAATCGAAAAACTGGGTGTAATTTGTCTAAGAGGATGCGAAGGGC 18
```

RESULT 2

```
AA606654/c
LOCUS       AA606654      464 bp    mRNA            EST       30-SEP-1997
DEFINITION  vm93a06.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
            clone IMAGE:1005778 5' similar to gb:L33715 Mus musculus Fau gene,
            complete cds (MOUSE);, mRNA sequence.
ACCESSION   AA606654
VERSION     AA606654.1  GI:2455588
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:693296.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:569994
            Putative full length read
            vector to vector length is 491.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Mus musculus"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1005778"
                     /clone_lib="Knowles Solter mouse blastocyst B1"
                     /tissue_type="blastocyst"
                     /dev_stage="embryo (pre-implantation)"
                     /lab_host="DH10B"
                     /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally from mRNA prepared
                     from 800 blastocysts.   Primer: SalI(dT):
                     5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs  were
                     cloned into the NotI/SalI sites of a pSPORT vector (Life
                     Technologies).  Two different size selections:  B1 (larger
                      inserts) and B3."
BASE COUNT      114 a    131 c    141 g     78 t
ORIGIN


  Query Match             38.6%;  Score 26.6;  DB 36;  Length 464;
  Best Local Similarity   66.7%;  Pred. No. 76;
  Matches   38;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCC 57
          |||  || ||||| |||       | |||| ||||  |||||  | ||| |||||||
Db    407 GGCACAACATTGACAAAGCGCCGGTTTTTACTGCATTCGGCCTTCTTGGCCCCGGCC 351


RESULT   3
AQ274419
LOCUS       AQ274419      692 bp    DNA             GSS       03-NOV-1998
DEFINITION  mgxb0015M12r CUGI Rice Blast BAC Library Pyricularia grisea genomic
            clone mgxb0015M12r, genomic survey sequence.
ACCESSION   AQ274419
VERSION     AQ274419.1  GI:3827734
KEYWORDS    GSS.
SOURCE      Pyricularia grisea.
  ORGANISM  Pyricularia grisea
            Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Pyriculariа.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
            Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
            Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson Universiy, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 390.
FEATURES             Location/Qualifiers
     source          1. .692
                     /organism="Pyricularia grisea"
                     /strain="70-15"
                     /db_xref="taxon:89476"
                     /clone="mgxb0015M12r"
                     /clone_lib="CUGI Rice Blast BAC Library"
                     /tissue_type="Protoplasts"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
                     Rice blast is one of the most devestating fungal diseases
                     of rice  world wide.  It is a filamentous ascomycete with
                     a haploid genome (n=7)  of approximately 40 Mbp.  Rice
                     blast is an important model fungal pathogen  for studying
                     numerous aspects of the fungal-host interaction.   In
                     order to facilitate genome wide analysis, a BAC library
                     containing  9216 clones with an average insert size of 130
                     kbp was constructed.   This library represents greater
                     than 25X genome coverage.  High density  colony filters
                     are available upon request."
BASE COUNT      145 a    169 c    210 g    167 t      1 others
ORIGIN


  Query Match             37.4%;  Score 25.8;  DB 105;  Length 692;
  Best Local Similarity   67.9%;  Pred. No. 1.3e+02;
  Matches   36;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      8 AATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
            |||  ||    |||| |||  ||  ||| ||||||| ||| ||| |   || |||
Db     59 AATGGACCGCCTAGGTCTTCTTCTTGGACGCGGCGGCCTTGGCCTTTGCAGCC 111


RESULT   4
CNS0164L
LOCUS       CNS0164L      1201 bp    DNA             GSS       26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15C21 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106287
VERSION     AL106287.1  GI:5621177
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota: Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
```

library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

```
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"
                     /clone_lib="DrosBAC"
                     /clone="BACN15C21"
                     /note="end : SP6"
BASE COUNT      304 a    204 c    230 g    275 t    188 others
ORIGIN


  Query Match             37.1%;  Score 25.6;  DB 83;  Length 1201;
  Best Local Similarity   32.8%;  Pred. No. 1.4e+02;
  Matches   20;  Conservative   26;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAAC 63
           ||| ||   ||||:::: :::::::|:|| : :::|::|::: :|  ||:  | ::|  :
Db     847 CAATAAMAMAAAAHHWWWWWHWWWWTWTTYSCSSSGSSCSSVGKGAACCSCCCNSSCSCS 906

Qy      64 G 64
           |
Db     907 G 907


RESULT   5
AQ406794
LOCUS       AQ406794      492 bp    DNA             GSS       17-MAR-1999
DEFINITION  HS_5103_B2_G10_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=679 Col=20 Row=N, genomic survey sequence.
ACCESSION   AQ406794
VERSION     AQ406794.1  GI:4429416
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 679  row: N  column: 20
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 492.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=679 Col=20 Row=N"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
BASE COUNT      170 a     96 c     93 g    131 t      2 others
ORIGIN


  Query Match             36.2%;  Score 25;  DB 106;  Length 492;
  Best Local Similarity   61.5%;  Pred. No. 2.4e+02;
  Matches   40;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
           || || | | || | |   ||| ||   |  || ||  ||| | |||| |||| |||  |
Db     239 GGACATATACTGGATATTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 298

Qy      61 AACGG 65
           || ||
Db     299 AATGG 303


RESULT   6
AQ212765
LOCUS       AQ212765      507 bp    DNA             GSS       18-SEP-1998
DEFINITION  HS_3115_B2_H11_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3115 Col=22 Row=P, genomic survey
            sequence.
ACCESSION   AQ212765
VERSION     AQ212765.1  GI:3623966
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     On Sep 10, 1998 this sequence version replaced gi:3553850.
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3115  row: P  column: 22
            Class: BAC ends
            High quality sequence stop: 507.
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3115 Col=22 Row=P"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      142 a    116 c    112 g    137 t
ORIGIN


  Query Match             36.2%;  Score 25;  DB 99;  Length 507;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+02;
  Matches   34;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       2 GCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGG 50
           ||||| |  || |||  || |  ||| ||||||   ||| |||||| ||
Db     154 GCCAATATCTGTAAAGATATACTTATATATTGCTATCGGTCGCCATAGG 202
```

```
RESULT   7
AQ699107
LOCUS       AQ699107      516 bp    DNA             GSS       06-JUL-1999
DEFINITION  HS_5560_B2_E01_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1136 Col=2 Row=J, genomic survey sequence.
ACCESSION   AQ699107
VERSION     AQ699107.1  GI:5389355
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     On Sep 10, 1998 this sequence version replaced gi:3555646.
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1136  row: J  column: 2
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 516.
FEATURES              Location/Qualifiers
     source           1. .516
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="Plate=1136 Col=2 Row=J"
                      /clone_lib="RPCI-11 Human Male BAC Library"
                      /sex="male"
                      /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
BASE COUNT      173 a    103 c    107 g    127 t      6 others
ORIGIN


  Query Match             36.2%;  Score 25;  DB 85;  Length 516;
  Best Local Similarity   61.5%;  Pred. No. 2.4e+02;
  Matches   40;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
           || || | | || | |   ||| ||   |  || ||  ||| | |||| |||| |||  |
Db     236 GGACATATACTGGATATTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 295

Qy      61 AACGG 65
           || ||
Db     296 AATGG 300


RESULT   8
AQ415486
LOCUS       AQ415486      603 bp    DNA             GSS       23-MAR-1999
DEFINITION  RPCI-11-177O19.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-177O19, genomic survey sequence.
ACCESSION   AQ415486
VERSION     AQ415486.1  GI:4474455
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .603
                      /organism="Homo sapiens"
                      /db_xref="GDB:7567938"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-177O19"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT      219 a    116 c    120 g    148 t
ORIGIN


  Query Match             36.2%;  Score 25;  DB 106;  Length 603;
  Best Local Similarity   61.5%;  Pred. No. 2.4e+02;
  Matches   40;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
           || || | | || |.|   ||| ||   |  || ||  ||| | |||| |||| |||  |
Db     237 GGACATATACTGGATATTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 296

Qy      61 AACGG 65
           || ||
Db     297 AATGG 301


RESULT   9
AQ384188
LOCUS       AQ384188      658 bp    DNA             GSS       21-MAY-1999
DEFINITION  RPCI11-122P9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-122P9,
            genomic survey sequence.
ACCESSION   AQ384188
VERSION     AQ384188.1  GI:4355211
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-122P9.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
```

```
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .658
                      /organism="Homo sapiens"
                      /db_xref="GDB:7546832"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-122P9"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT      235 a    128 c    126 g    169 t
ORIGIN


  Query Match             36.2%;  Score 25;  DB 106;  Length 658;
  Best Local Similarity   61.5%;  Pred. No. 2.4e+02;
  Matches   40;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
           || || | | || | |   ||| ||   |  || ||  ||| | |||| |||| |||  |
Db     225 GGACATATACTGGATATTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 284

Qy      61 AACGG 65
           || ||
Db     285 AATGG 289


RESULT  10
AW265066/c
LOCUS       AW265066      426 bp    mRNA            EST       28-DEC-1999
DEFINITION  xq60g03.x1 NCI_CGAP_Co22 Homo sapiens cDNA clone IMAGE:2755060 3'
            similar to SW:E411_ADE02 P03241 PROBABLE EARLY E4 11 KD PROTEIN.
            [1] ;, mRNA sequence.
ACCESSION   AW265066
VERSION     AW265066.1  GI:6641882
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jul 7, 1999 this sequence version replaced gi:5406301.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            unknown library type
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 425.
FEATURES              Location/Qualifiers
     source           1. .426
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2755060"
                      /clone_lib="NCI_CGAP_Co22"
                      /tissue_type="colonic adenocarcinoma"
                      /lab_host="DH10B"
                      /note="Organ: colon; Vector: pAMP10; cDNA made by oligo-dT
                      priming.  Non-directionally cloned into the UDG sites of
                      pAMP10.  Size-selected on agarose gel, average insert
                      size 500 bp.  Primary library; non-amplified.   cDNA
                      Library Preparation: David B. Krizman, Ph.D (NCI).
                      Reference:  Krizman et al. (1996) Cancer Research
                      56:5380-5383."
BASE COUNT      103 a     80 c    118 g    124 t      1 others
ORIGIN


  Query Match             35.9%;  Score 24.8;  DB 79;  Length 426;
  Best Local Similarity   63.3%;  Pred. No. 2.8e+02;
  Matches   38;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAA 62
           ||||||  | ||||   |||||||||| | || |||||  | ||   ||     |  | |||
Db     257 CCAAAACCTCAAAATGAAGATCTATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCAA 198


RESULT  11
AW276836/c
LOCUS       AW276836      630 bp    mRNA            EST       03-JAN-2000
DEFINITION  xp66g11.x1 NCI_CGAP_Ov39 Homo sapiens cDNA clone IMAGE:2745380 3'
            similar to SW:E413_ADE02 P03240 PROBABLE EARLY E4 13 KD PROTEIN. ;,
            mRNA sequence.
ACCESSION   AW276836
VERSION     AW276836.1  GI:6663866
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Nov 22, 1999 this sequence version replaced gi:6462296.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Andrew Berchuck M.D., John Gillespie M.D.,
            Michael R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 422.
FEATURES              Location/Qualifiers
     source           1. .630
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2745380"
                      /clone_lib="NCI_CGAP_Ov39"
                      /sex="female"
                      /tissue_type="papillary serous ovarian metastasis"
                      /lab_host="DH10B"
                      /note="Organ: ovary; Vector: pAMP10; cDNA made by oligo-dT
                      priming.  Non-directionally cloned into the UDG sites of
                      pAMP10.  Size-selected on agarose gel, average insert
                      size 500 bp.  Primary library; non-amplified.   cDNA
                      Library Preparation: David B. Krizman, Ph.D (NCI).
                      Reference:  Krizman et al. (1996) Cancer Research
```

```
                    56:5380-5383."
BASE COUNT      154 a    134 c    173 g    168 t      1 others
ORIGIN


  Query Match             35.9%;  Score 24.8;  DB 79;  Length 630;
  Best Local Similarity   63.3%;  Pred. No. 2.7e+02;
  Matches   38;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAA 62
          ||||||  | ||||   |||||||||| | || |||||  | ||   ||    |  | |||
Db    290 CCAAAACCTCAAAATGAAGATCTATTAAGTGAACGCGCTCCCCTCCGGTGGCGTGGTCAA 231


RESULT  12
C84505
LOCUS       C84505         379 bp    mRNA            EST       26-MAR-1999
DEFINITION  C84505 osteoclast subtracted library Oryctolagus cuniculus cDNA,
            mRNA sequence.
ACCESSION   C84505
VERSION     C84505.1  GI:4527765
KEYWORDS    EST.
SOURCE      Oryctolagus cuniculus.
  ORGANISM  Oryctolagus cuniculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Kobori,M., Ikeda,Y., Nara,H., Kato,M., Kumegawa,M., Nojima,H. and
            Kawashima,H.
  TITLE     Large scale isolation of osteoclast-specific genes by an improved
            method involving the preparation of a subtracted cDNA library
  JOURNAL   Genes Cells 3 (7), 459-475 (1998)
  MEDLINE   98424349
COMMENT     On Jun 5, 1998 this sequence version replaced gi:3189094.
            Contact: Kobori M
            Molecular Medicine Laboratories
            Institute for Drug Discovery Research,Yamanouchi Pharmaceutica
            21, Miyukigaoka, Tsukuba, Ibaraki 305, Japan
            Email: kobori@yamanouchi.co.jp
            PROJECT = 'OSG'.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Oryctolagus cuniculus"
                     /db_xref="taxon:9986"
                     /clone_lib="osteoclast subtracted library"
                     /tissue_type="long bone"
                     /cell_type="osteoclast"
                     /cell_line="primary"
                     /dev_stage="5 day-old"
BASE COUNT       72 a    125 c    108 g     58 t     16 others
ORIGIN


  Query Match             35.7%;  Score 24.6;  DB 48;  Length 379;
  Best Local Similarity   60.9%;  Pred. No. 3.3e+02;
  Matches   39;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCC 60
          || | |||  | | || || || | ||||  ||  |  |||  ||   |  |||||| ||
Db     33 GGACCAAACCTCAGAACCTGGACCCATTTGNTGACCTGGGCGACCTCAGCTCCGGCCTCC 92

Qy     61 AACG 64
          || |
Db     93 AAGG 96


RESULT  13
AW354184
LOCUS       AW354184       194 bp    mRNA            EST       02-FEB-2000
DEFINITION  32589 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW354184
VERSION     AW354184.1  GI:6853174
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 194)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A. and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     On May 18, 1998 this sequence version replaced gi:3138121.
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 13  row: A  column: 24
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .194
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       33 a     50 c     93 g     18 t
ORIGIN


  Query Match             35.4%;  Score 24.4;  DB 80;  Length 194;
  Best Local Similarity   63.8%;  Pred. No. 4.1e+02;
  Matches   37;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy     11 TGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAACGGCGG 68
          |||| | | || |  |    | | | || | |||| ||| |||||||| || ||  ||
Db     21 TGAAGAGCGAGGTGGAGGCCTCGGAGGCCGGCGCCCTGGACCCGGCCGGCAGCGCAGG 78


RESULT  14
AQ064961/c
LOCUS       AQ064961       355 bp    DNA             GSS       04-AUG-1998
DEFINITION  HS_2225_A1_G06_MF CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2225 Col=11 Row=M, genomic survey
            sequence.
ACCESSION   AQ064961
VERSION     AQ064961.1  GI:3380789
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
```

```
          Tel: (206) 616-3618
          Fax: (206) 616-3887
          Email: jwallace@u.washington.edu
          Sequence Tagged Connector
          Plate: 2225  row: M  column: 11
          Class: BAC ends
          High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2225 Col=11 Row=M"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      111 a     70 c     71 g    103 t
ORIGIN


  Query Match             35.4%;  Score 24.4;  DB 96;  Length 355;
  Best Local Similarity   73.8%;  Pred. No. 3.9e+02;
  Matches   31;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGC 44
           | || |||| | ||| || || || ||| ||||||| || ||| |
Db     197 CAAAGAATTAACAAAGTACATGTATCTATTGCAAGCAGCCCC 156


RESULT  15
C78547/c
LOCUS       C78547         578 bp    mRNA            EST       25-JUN-1998
DEFINITION  C78547 Mouse 3.5-dpc blastocyst cDNA Mus musculus cDNA clone
            J0051B03 3' similar to Mus musculus transcription factor like
            protein 4 TCFL4, mRNA sequence.
ACCESSION   C78547
VERSION     C78547.1  GI:2518877
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Ko,M.S.H., Kitchen,J.R., Wang,X., Wang,X., Threat,T.A., Sun,T.,
            Grahovac,M.J., Mason,S., Lim,M.K., Paonessa,P.D., Sauls,A.D. and
            Doi,H.
  TITLE     Systematic analyses of genes expressed in 3.5-dpc mouse blastocyst
            (The ERATO/Doi Project at Wayne State University)
  JOURNAL   Unpublished (1997)
COMMENT     On May 5, 1995 this sequence version replaced gi:797882.
            Contact: Hirofumi Doi
            Doi Bioasymmetry Project, ERATO
            Japan Science and Technology Corporation (JST)
            WBG Marive East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
            Email: hd@bioa.jst.go.jp.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="J0051B03"
                     /clone_lib="Mouse 3.5-dpc blastocyst cDNA"
                     /tissue_type="blastocyst"
                     /dev_stage="3.5-dpc"
BASE COUNT      168 a    128 c    152 g    128 t      2 others
ORIGIN


  Query Match             35.4%;  Score 24.4;  DB 36;  Length 578;
  Best Local Similarity   62.7%;  Pred. No. 3.7e+02;
  Matches   37;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGGCCCGGCCGCCAA 62
           || | |||||||||| | |||| | ||||| |  || |||    |  | |||  |||
Db     448 CATCAGTTGAAAAACCAACTCTACTGATTGCCCATGGAAGCCNGCAGAGCAGCCAACAA 390


Search completed: May 29, 2000, 21:13:10
Job time: 36509 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:33:53 ; Search time 1214.87 Seconds
(without alignments)
470.244 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GGCCAAAAATTGAAATTTTA.....GCCCGGCCGCCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		40.6	54.1	2164	16	VVH3FDPC1	X89856 V.virus DNA
2	c	40.6	54.1	2164	16	VVH3FDPC2	X89857 V.virus DNA
3		40	53.3	84	5	A59039	A59039 Sequence 27
4	c	40	53.3	86	5	A59040	A59040 Sequence 28
5		32.4	43.2	160821	54	AC012468	AC012468 Homo sapi
6	c	31.8	42.4	331068	41	AC009246	AC009246 Homo sapi
7	c	31	41.3	4376	12	MUSIL13A	L13028 Mouse inter
8		31	41.3	159500	12	AC005742	AC005742 Mus muscu
9		31	41.3	196382	55	AC020886	AC020886 Mus muscu
10		30.4	40.5	201746	41	AC006895	AC006895 Caenorhab
11	c	30.2	40.3	161449	40	AC005341	AC005341 Homo sapi
12	c	30.2	40.3	161986	45	AC021672	AC021672 Homo sapi
13		29.8	39.7	108061	8	AC009526	AC009526 Arabidops
14	c	29.8	39.7	141655	32	HS1093G12	AL121751 Homo sapi
15		29.8	39.7	218305	33	HSJ324O17	AL110115 Homo sapi
16	c	29.8	39.7	218305	33	HSJ324O17	AL110115 Homo sapi
17		29.6	39.5	62944	40	HSMHC3W5A	U89336 Homo sapien
18		29.6	39.5	183285	42	AC009833	AC009833 Homo sapi
19	c	29.6	39.5	275670	55	AC022328	AC022328 Mus muscu
20	c	29.4	39.2	35184	42	AC014624	AC014624 Drosophil
21	c	29.4	39.2	39573	35	CELK08D9	AF101319 Caenorhab
22		29.2	38.9	35974	34	CELC26F1	U53148 Caenorhabdi
23		29.2	38.9	134095	41	AC006915	AC006915 Caenorhab
24	c	29.2	38.9	142232	56	AC008443	AC008443 Homo sapi
25	c	29.2	38.9	179100	42	AC012641	AC012641 Homo sapi
26		29.2	38.9	227066	41	AC006877	AC006877 Caenorhab
27	c	29.2	38.9	298804	41	AC006911	AC006911 Caenorhab
28	c	29	38.7	31531	34	CEZC434	Z75714 Caenorhabdi
29	c	29	38.7	73461	56	AC010563	AC010563 Drosophil
30	c	29	38.7	110000	32	CEY106G6_5	Continuation (6 of
31		29	38.7	164197	44	AC016328	AC016328 Homo sapi
32		29	38.7	191258	45	AC008462	AC008462 Homo sapi
33		28.8	38.4	61450	10	HS15D7	AL031229 Human DNA
34	c	28.8	38.4	96308	40	AF190465	AF190465 Homo sapi
35		28.8	38.4	149259	43	AC013668	AC013668 Homo sapi
36		28.8	38.4	151595	54	AC015809	AC015809 Homo sapi
37		28.8	38.4	165093	41	AC005517	AC005517 Homo sapi
38	c	28.8	38.4	256172	41	AC005139	AC005139 Plasmodiu
39		28.6	38.1	2166	35	AF200327	AF200327 Plasmodiu
40		28.6	38.1	27885	34	CELZK488	AF022984 Caenorhab
41		28.6	38.1	30452	9	AB014077	AB014077 Homo sapi
42	c	28.6	38.1	38561	41	AC006604	AC006604 Caenorhab
43		28.6	38.1	43600	11	AC004209	AC004209 Homo sapi
44		28.6	38.1	100000	9	AP000514	AP000514 Homo sapi
45	c	28.6	38.1	101584	41	AC006793	AC006793 Caenorhab

ALIGNMENTS

```
RESULT   1
VVH3FDPC1
LOCUS       VVH3FDPC1    2164 bp    DNA             VRL       03-JAN-1996
DEFINITION  V.virus DNA for Hind III F fragment (clone VDPC1).
ACCESSION   X89856
VERSION     X89856.1  GI:927568
KEYWORDS    HindIII F fragment; hph gene; hygromycin phosphotransferase;
            promoter.
SOURCE      Vaccinia virus.
  ORGANISM  Vaccinia virus
            Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
            Orthopoxvirus.
REFERENCE   1  (bases 1 to 2164)
  AUTHORS   Pfleiderer,M., Falkner,F.G. and Dorner,F.
  TITLE     A novel vaccinia virus expression system allowing construction of
            recombinants without the need for selection markers, plasmids and
            bacterial hosts
  JOURNAL   J. Gen. Virol. 76 (Pt 12), 2957-2962 (1995)
  MEDLINE   96112180
REFERENCE   2  (bases 1 to 2164)
  AUTHORS   Pfleiderer,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-1995) M. Pfleiderer, Immuno AG, A. 2304
            Orth/Donau, Uferstrasse 15, AUSTRIA
COMMENT     Related sequence M35027.
FEATURES             Location/Qualifiers
     source          1. .2164
                     /organism="Vaccinia virus"
                     /strain="WR"
                     /specific_host="mammalians"
                     /db_xref="taxon:10245"
                     /lab_host="mammalian cell lines"
                     /cell_line="recombinant Vaccinia virus"
                     /clone="VDPC 1"
                     /map="45478 nt number"
     misc_feature    1. .485
                     /note="Hind III F fragment"
     misc_feature    479. .1662
                     /note="NOT I gene cassette"
     misc_feature    479
                     /note="NOT I cleavage site corresponding to the unique NOT
                     I site in the Vv genome at position 45478 (strain
                     copenhagen)"
     promoter        486. .543
                     /note="strong synthetic early/late promoter"
     misc_feature    536
                     /note="unique Sfi I cleavage site"
     promoter        544. .594
                     /note="weak synthetic early/late promoter"
     gene            595. .1620
                     /gene="hph"
     CDS             595. .1620
                     /gene="hph"
                     /codon_start=1
                     /product="hygromycin phosphotransferase"
                     /protein_id="CAA61952.1"
                     /db_xref="GI:953210"
                     /db_xref="SWISS-PROT:P00557"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE"
     misc_feature    1635
                     /note="Rsr II cleavage site"
     misc_feature    1642
                     /note="NOT I cleavage site"
     misc_feature    1643. .2164
                     /note="Hind III F fragment"

BASE COUNT      586 a    488 c    475 g    615 t
ORIGIN
```

Query Match 54.1%; Score 40.6; DB 16; Length 2164;
Best Local Similarity 73.2%; Pred. No. 0.038;
Matches 52; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCG 64
           |||||||||||||||||||||||||||||||||||||   | |   |  |||||    |
Db     488 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCGAAGAGGCCCATATG 547

Qy      65 CCAACGGCGGA 75
             || |  | |
Db     548 TAAAAGTTGAA 558
```

```
RESULT   2
VVH3FDPC2/c
LOCUS       VVH3FDPC2    2164 bp    DNA             VRL       03-JAN-1996
DEFINITION  V.virus DNA for Hind III F fragment (clone vDPC2).
ACCESSION   X89857
VERSION     X89857.1  GI:927569
KEYWORDS    HindIII F fragment; hph gene; hygromycin phosphotransferase;
            promoter.
SOURCE      Vaccinia virus.
  ORGANISM  Vaccinia virus
            Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
            Orthopoxvirus.
REFERENCE   1  (bases 1 to 2164)
  AUTHORS   Pfleiderer,M., Falkner,F.G. and Dorner,F.
  TITLE     A novel vaccinia virus expression system allowing construction of
            recombinants without the need for selection markers, plasmids and
            bacterial hosts
  JOURNAL   J. Gen. Virol. 76 (Pt 12), 2957-2962 (1995)
  MEDLINE   96112180
REFERENCE   2  (bases 1 to 2164)
  AUTHORS   Pfleiderer,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-1995) M. Pfleiderer, Immuno AG, A. 2304
            Orth/Donau, Uferstrasse 15, AUSTRIA
COMMENT     Related sequence M35027.
FEATURES             Location/Qualifiers
     source          1. .2164
                     /organism="Vaccinia virus"
                     /strain="WR"
                     /specific_host="mammalians"
                     /db_xref="taxon:10245"
                     /lab_host="Mammalian cell lines"
                     /cell_line="recombinant Vaccinia virus"
                     /clone="VDPC 2"
                     /map="45478 nt number"
     misc_feature    1. .485
                     /note="Hind III F fragment"
     misc_feature    479
                     /note="NOT cleavage site corresponding to the unique NOT I
                     site in the Vv genome at position 45478 (strain
                     copenhagen)"
     misc_feature    complement(479. .1642)
                     /note="NOT I gene cassette"
     misc_feature    487
                     /note="Rsr II cleavage site"
     gene            complement(506. .1531)
                     /gene="hph"
     CDS             complement(506. .1531)
                     /gene="hph"
                     /codon_start=1
                     /product="hygromycin phosphotransferase"
                     /protein_id="CAA61953.1"
                     /db_xref="GI:953211"
                     /db_xref="SWISS-PROT:P00557"
                     /translation="MKKPELTATSVEKFLIEKFDSVSDLMQLSEGEESRAFSFDVGGR
                     GYVLRVNSCADGFYKDRYVYRHFASAALPIPEVLDIGEFSESLTYCISRRAQGVTLQD,
```

```
                     LPETELPAVLQPVAEAMDAIAAADLSQTSGFGPFGPQGIGQYTTWRDFICAIADPHVY
                     HWQTVMDDTVSASVAQALDELMLWAEDCPEVRHLVHADFGSNNVLTDNGRITAVIDWS
                     EAMFGDSQYEVANIFFWRPWLACMEQQTRYFERRHPELAGSPRLRAYMLRIGLDQLYQ
                     SLVDGNFDDAAWAQGRCDAIVRSGAGTVGRTQIARRSAAVWTDGCVEVLADSGNRRPS
                     TRPRAKE"
     promoter        complement(1532. .1585)
                     /note="weak synthetic early/late promoter"
     misc_feature    1592
                     /note="Sfi cleavage site"
     promoter        complement(1598. .1639)
                     /note="strong synthetic early/late promoter"
     misc_feature    1642
                     /note="NOT I cleavage site"
     misc_feature    1643. .2164
                     /note="Hind III F fragment"
BASE COUNT      595 a    524 c    439 g    606 t
ORIGIN


  Query Match             54.1%;  Score 40.6;  DB 16;  Length 2164;
  Best Local Similarity   73.2%;  Pred. No. 0.038;
  Matches   52;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCG 64
          |||||||||||||||||||||||||||||||||||||||   | |   |  |||||    |
Db   1638 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCGAAGAGGCCCATATG 1579

Qy     65 CCAACGGCGGA 75
            || |  | |
Db   1578 TAAAAGTTGAA 1568


RESULT   3
A59039
LOCUS       A59039         84 bp    DNA             PAT       06-MAR-1998
DEFINITION  Sequence 27 from Patent EP0753581.
ACCESSION   A59039
VERSION     A59039.1  GI:3714474
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Scheiflinger,F.D., Antoine,G.D., Falkner, Falko-Guenter,D.,
            Dorner,F.P. and Eibl,J.D.
  TITLE     Improved recombinant eukaryotic cytoplasmic viruses, method for
            their production and their use as vaccines
  JOURNAL   Patent: EP 0753581-A 27 15-JAN-1997;
            IMMUNO AG (AT)
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       28 a      8 c     10 g     38 t
ORIGIN


  Query Match             53.3%;  Score 40;  DB 5;  Length 84;
  Best Local Similarity   82.1%;  Pred. No. 0.038;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
          | ||||||||||||||||||||||||||||||||||   | |  |   ||||
Db      9 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 64


RESULT   4
A59040/c
LOCUS       A59040         86 bp    DNA             PAT       06-MAR-1998
DEFINITION  Sequence 28 from Patent EP0753581.
ACCESSION   A59040
VERSION     A59040.1  GI:3714475
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Scheiflinger,F.D., Antoine,G.D., Falkner, Falko-Guenter,D.,
            Dorner,F.P. and Eibl,J.D.
  TITLE     Improved recombinant eukaryotic cytoplasmic viruses, method for
            their production and their use as vaccines
  JOURNAL   Patent: EP 0753581-A 28 15-JAN-1997;
            IMMUNO AG (AT)
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       38 a     11 c      9 g     28 t
ORIGIN


  Query Match             53.3%;  Score 40;  DB 5;  Length 86;
  Best Local Similarity   82.1%;  Pred. No. 0.038;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
          | ||||||||||||||||||||||||||||||||||   | |  |   ||||
Db     80 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 25


RESULT   5
AC012468
LOCUS       AC012468   160821 bp    DNA             HTG       03-FEB-2000
DEFINITION  Homo sapiens clone RP11-179H18, WORKING DRAFT SEQUENCE, 29
            unordered pieces.
ACCESSION   AC012468
VERSION     AC012468.5  GI:6693085
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160821)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160821)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-OCT-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jan 14, 2000 this sequence version replaced gi:6682197.

                        -------------- Genome Center
                            Center: Genome Therapeutics Corperation
                            Center code: GTC
                            Web site: http://www.genomecorp.com/
                            Contact: gtc-seqcenter@genomecorp.com
                        -------------- Project Information
                            Center project name: hg050
                        -------------- Summary Statistics
                            Sequencing vector: N/A
                            Chemistry: Dye-terminator Big Dye; 100% of reads
                            Assembly program: Phrap; version 990315
                            Consensus quality: 130458 bases at least Q40
                            Consensus quality: 146152 bases at least Q30
                            Consensus quality: 149165 bases at least Q20
                            Insert size: 160821; sum-of-contigs
                            Quality coverage: 4.1x in Q20 bases; sum-of-contigs
                        --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 29 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
```

```
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1162: contig of 1162 bp in length
            *                    gap of unknown length
            *     1163     2431: contig of 1269 bp in length
            *                    gap of unknown length
            *     2432     3867: contig of 1436 bp in length
            *                    gap of unknown length
            *     3868     5091: contig of 1224 bp in length
            *                    gap of unknown length
            *     5092     6952: contig of 1861 bp in length
            *                    gap of unknown length
            *     6953     8125: contig of 1173 bp in length
            *                    gap of unknown length
            *     8126     9518: contig of 1393 bp in length
            *                    gap of unknown length
            *     9519    10648: contig of 1130 bp in length
            *                    gap of unknown length
            *    10649    11759: contig of 1111 bp in length
            *                    gap of unknown length
            *    11760    13202: contig of 1443 bp in length
            *                    gap of unknown length
            *    13203    15433: contig of 2231 bp in length
            *                    gap of unknown length
            *    15434    17627: contig of 2194 bp in length
            *                    gap of unknown length
            *    17628    19656: contig of 2029 bp in length
            *                    gap of unknown length
            *    19657    22735: contig of 3079 bp in length
            *                    gap of unknown length
            *    22736    25119: contig of 2384 bp in length
            *                    gap of unknown length
            *    25120    27919: contig of 2800 bp in length
            *                    gap of unknown length
            *    27920    31649: contig of 3730 bp in length
            *                    gap of unknown length
            *    31650    37383: contig of 5734 bp in length
            *                    gap of unknown length
            *    37384    41600: contig of 4217 bp in length
            *                    gap of unknown length
            *    41601    44764: contig of 3164 bp in length
            *                    gap of unknown length
            *    44765    50654: contig of 5890 bp in length
            *                    gap of unknown length
            *    50655    58608: contig of 7954 bp in length
            *                    gap of unknown length
            *    58609    66645: contig of 8037 bp in length
            *                    gap of unknown length
            *    66646    76568: contig of 9923 bp in length
            *                    gap of unknown length
            *    76569    86608: contig of 10040 bp in length
            *                    gap of unknown length
            *    86609   101468: contig of 14860 bp in length
            *                    gap of unknown length
            *   101469   118027: contig of 16559 bp in length
            *                    gap of unknown length
            *   118028   140795: contig of 22768 bp in length
            *                    gap of unknown length
            *   140796   160821: contig of 20026 bp in length.
FEATURES             Location/Qualifiers
     source          1. .160821
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-179H18"
                     /clone_lib="RPCI-11"
BASE COUNT    43392 a  36669 c  36666 g  44015 t     79 others
ORIGIN


  Query Match             43.2%;  Score 32.4;  DB 54;  Length 160821;
  Best Local Similarity   71.2%;  Pred. No. 15;
  Matches   42;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy       6 AAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCG 64
           | || | ||| ||| |||||||||||| ||        | ||| ||| ||||||||| ||
Db   27681 ATAACTTAAACTTTTTTTTTTTTTTTTTTGAGCNGGGGGGGGCAGCCGTGGGCCCGGGCG 27739


RESULT   6
AC009246/c
LOCUS       AC009246   331068 bp    DNA             HTG       03-SEP-1999
DEFINITION  Homo sapiens, *** SEQUENCING IN PROGRESS ***, 96 unordered pieces.
ACCESSION   AC009246
VERSION     AC009246.7  GI:5819117
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 331068)
  AUTHORS   Muzny,D.M., Adams,C., Arenson,A., Bailey,M., Barbaria,J.,
            Blankenburg,K., Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C.,
            Bunac,C., Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z.,
            Cox,C., David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lu,J., Lucier,R.,
            Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M., Morris,S.,
            Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G.,
            Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M.,
            Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E.,
            Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R., Tabor,P.,
            Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,
            Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 331068)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 2, 1999 this sequence version replaced gi:5787982.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 96 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    74449: contig of 74449 bp in length
            *    74450    74469: gap of unknown length
            *    74470    91427: contig of 16958 bp in length
            *    91428    91447: gap of unknown length
            *    91448   106480: contig of 15033 bp in length
            *   106481   106500: gap of unknown length
            *   106501   119379: contig of 12879 bp in length
            *   119380   119399: gap of unknown length
            *   119400   130144: contig of 10745 bp in length
            *   130145   130164: gap of unknown length
            *   130165   140199: contig of 10035 bp in length
            *   140200   140219: gap of unknown length
            *   140220   149764: contig of 9545 bp in length
            *   149765   149784: gap of unknown length
            *   149785   157914: contig of 8130 bp in length
            *   157915   157934: gap of unknown length
            *   157935   164329: contig of 6395 bp in length
            *   164330   164349: gap of unknown length
            *   164350   170609: contig of 6260 bp in length
            *   170610   170629: gap of unknown length
```

```
*   170630   176330: contig of 5701 bp in length
*   176331   176350: gap of unknown length
*   176351   181979: contig of 5629 bp in length
*   181980   181999: gap of unknown length
*   182000   187233: contig of 5234 bp in length
*   187234   187253: gap of unknown length
*   187254   192438: contig of 5185 bp in length
*   192439   192458: gap of unknown length
*   192459   196875: contig of 4417 bp in length
*   196876   196895: gap of unknown length
*   196896   201161: contig of 4266 bp in length
*   201162   201181: gap of unknown length
*   201182   205310: contig of 4129 bp in length
*   205311   205330: gap of unknown length
*   205331   209336: contig of 4006 bp in length
*   209337   209356: gap of unknown length
*   209357   213217: contig of 3861 bp in length
*   213218   213237: gap of unknown length
*   213238   217097: contig of 3860 bp in length
*   217098   217117: gap of unknown length
*   217118   220881: contig of 3764 bp in length
*   220882   220901: gap of unknown length
*   220902   224582: contig of 3681 bp in length
*   224583   224602: gap of unknown length
*   224603   228191: contig of 3589 bp in length
*   228192   228211: gap of unknown length
*   228212   231355: contig of 3144 bp in length
*   231356   231375: gap of unknown length
*   231376   234318: contig of 2943 bp in length
*   234319   234338: gap of unknown length
*   234339   237245: contig of 2907 bp in length
*   237246   237265: gap of unknown length
*   237266   240108: contig of 2843 bp in length
*   240109   240128: gap of unknown length
*   240129   242952: contig of 2824 bp in length
*   242953   242972: gap of unknown length
*   242973   245759: contig of 2787 bp in length
*   245760   245779: gap of unknown length
*   245780   248484: contig of 2705 bp in length
*   248485   248504: gap of unknown length
*   248505   250933: contig of 2429 bp in length
*   250934   250953: gap of unknown length
*   250954   253206: contig of 2253 bp in length
*   253207   253226: gap of unknown length
*   253227   255474: contig of 2248 bp in length
*   255475   255494: gap of unknown length
*   255495   257728: contig of 2234 bp in length
*   257729   257748: gap of unknown length
*   257749   259941: contig of 2193 bp in length
*   259942   259961: gap of unknown length
*   259962   262038: contig of 2077 bp in length
*   262039   262058: gap of unknown length
*   262059   264107: contig of 2049 bp in length
*   264108   264127: gap of unknown length
*   264128   266112: contig of 1985 bp in length
*   266113   266132: gap of unknown length
*   266133   267955: contig of 1823 bp in length
*   267956   267975: gap of unknown length
*   267976   269749: contig of 1774 bp in length
*   269750   269769: gap of unknown length
*   269770   271539: contig of 1770 bp in length
*   271540   271559: gap of unknown length
*   271560   273287: contig of 1728 bp in length
*   273288   273307: gap of unknown length
*   273308   275008: contig of 1701 bp in length
*   275009   275028: gap of unknown length
*   275029   276712: contig of 1684 bp in length
*   276713   276732: gap of unknown length
*   276733   278265: contig of 1533 bp in length
*   278266   278285: gap of unknown length
*   278286   279784: contig of 1499 bp in length
*   279785   279804: gap of unknown length
*   279805   281282: contig of 1478 bp in length
*   281283   281302: gap of unknown length
*   281303   282707: contig of 1405 bp in length
*   282708   282727: gap of unknown length
*   282728   284093: contig of 1366 bp in length
*   284094   284113: gap of unknown length
*   284114   285441: contig of 1328 bp in length
*   285442   285461: gap of unknown length
*   285462   286777: contig of 1316 bp in length
*   286778   286797: gap of unknown length
*   286798   288075: contig of 1278 bp in length
*   288076   288095: gap of unknown length
*   288096   289370: contig of 1275 bp in length
*   289371   289390: gap of unknown length
*   289391   290655: contig of 1265 bp in length
*   290656   290675: gap of unknown length
*   290676   291920: contig of 1245 bp in length
*   291921   291940: gap of unknown length
*   291941   293182: contig of 1242 bp in length
*   293183   293202: gap of unknown length
*   293203   294432: contig of 1230 bp in length
*   294433   294452: gap of unknown length
*   294453   295681: contig of 1229 bp in length
*   295682   295701: gap of unknown length
*   295702   296919: contig of 1218 bp in length
*   296920   296939: gap of unknown length
*   296940   298153: contig of 1214 bp in length
*   298154   298173: gap of unknown length
*   298174   299376: contig of 1203 bp in length
*   299377   299396: gap of unknown length
*   299397   300537: contig of 1141 bp in length
*   300538   300557: gap of unknown length
*   300558   301692: contig of 1135 bp in length
*   301693   301712: gap of unknown length
*   301713   302794: contig of 1082 bp in length
*   302795   302814: gap of unknown length
*   302815   303887: contig of 1073 bp in length
*   303888   303907: gap of unknown length
*   303908   304974: contig of 1067 bp in length
*   304975   304994: gap of unknown length
*   304995   306059: contig of 1065 bp in length
*   306060   306079: gap of unknown length
*   306080   307072: contig of 993 bp in length
*   307073   307092: gap of unknown length
*   307093   308077: contig of 985 bp in length
*   308078   308097: gap of unknown length
*   308098   309080: contig of 983 bp in length
*   309081   309100: gap of unknown length
*   309101   310081: contig of 981 bp in length
*   310082   310101: gap of unknown length
*   310102   311057: contig of 956 bp in length
*   311058   311077: gap of unknown length
*   311078   312000: contig of 923 bp in length
*   312001   312020: gap of unknown length
*   312021   312939: contig of 919 bp in length
*   312940   312959: gap of unknown length
*   312960   313863: contig of 904 bp in length
*   313864   313883: gap of unknown length
*   313884   314773: contig of 890 bp in length
*   314774   314793: gap of unknown length
*   314794   315682: contig of 889 bp in length
*   315683   315702: gap of unknown length
*   315703   316578: contig of 876 bp in length

  Query Match            42.4%;  Score 31.8;  DB 41;  Length 331068;
  Best Local Similarity  71.2%;  Pred. No. 25;
  Matches  42;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGCGG 74
           |||| ||||||||||||  || | |||| | |||   ||||| |||  ||  ||| | |
Db  64735 TTTTTTTTTTTTTTTTTTTAAAAAAAAGGGACCGGGCTGGGCGCGGTGGCTCACGCCTG 64677

RESULT   7
```

```
MUSIL13A/c
LOCUS       MUSIL13A     4376 bp    DNA             ROD       27-JUL-1993
DEFINITION  Mouse interleukin-13 gene sequence with four exons.
ACCESSION   L13028
VERSION     L13028.1  GI:309397
KEYWORDS    interleukin 13.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) adult liver
            DNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4376)
  AUTHORS   McKenzie,A.N.J., Li,X., Largaespada,D.A., Sato,A., Kaneda,A.,
            Zurawski,S.M., Doyle,E.L., Milatovich,A., Francke,U.,
            Copeland,N.G., Jenkins,N.A. and Zurawski,G.
  TITLE     Structural comparison and chromosomal localization of the human and
            mouse IL-13 genes
  JOURNAL   J. Immunol. 150, 5436-5444 (1993)
  MEDLINE   93294299
FEATURES             Location/Qualifiers
     source          1. .4376
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /dev_stage="adult"
                     /germline
                     /tissue_type="liver"
     exon            <726. .934
     exon            2193. .2246
                     /number=2
     exon            2822. .2927
                     /number=3
     exon            3240. .>4073
                     /number=4
BASE COUNT     1062 a    1122 c    1146 g    1046 t
ORIGIN
```

```
  Query Match           41.3%;  Score 31;  DB 12;  Length 4376;
  Best Local Similarity  78.7%;  Pred. No. 25;
  Matches   37;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      10 TTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGG 56
           ||||| |||| ||||| |||||||||| ||  || ||||||  |  ||
Db     622 TTGAATTTTTTTTTTTCTTTTTTGGAAAATCCAGTGGCCGCTAAAGG 576
```

```
RESULT   8
AC005742
LOCUS       AC005742   159500 bp    DNA             ROD       01-OCT-1998
DEFINITION  Mus musculus chromosome 11, BAC clone 111-181 (LBNL M01), complete
            sequence.
ACCESSION   AC005742
VERSION     AC005742.1  GI:3687208
KEYWORDS    HTG.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 159500)
  AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
  TITLE     Sequencing of mouse chromosome 11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 159500)
  AUTHORS   Ricke,D.O.
  TITLE     Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 159500)
  AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1998) Human Genome Center, DOE Joint Genome
            Institute, Lawrence Berkeley National Laboratory, MS 74-157,
            Berkeley, CA 94720, U.S.A.
COMMENT     Sequence submitted by:
            DOE Joint Genome Institute.
FEATURES             Location/Qualifiers
     source          1. .159500
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="11"
                     /clone="111-181"
                     /chromosome="11"
                     /note="LBNL M01"
     repeat_region   complement(616. .691)
                     /rpt_family="Alu"
     repeat_region   2368. .2409
                     /note="(CT)21"
                     /rpt_type=tandem
                     /rpt_unit=CT
     repeat_region   2522. .2559
                     /note="(GT)19"
                     /rpt_type=tandem
                     /rpt_unit=GT
     repeat_region   3628. .3659
                     /note="(AC)16"
                     /rpt_type=tandem
                     /rpt_unit=AC
     misc_feature    5290. .5410
                     /note="GRAIL 2 excellent exon, frame 0"
     misc_feature    7727. .7953
                     /note="GRAIL 2 excellent exon, frame 0"
     misc_feature    8243. .8429
                     /note="GRAIL 2 excellent exon, frame 2"
     misc_feature    9242. .9407
                     /note="GRAIL 2 excellent exon, frame 1"
     misc_feature    10448. .10616
                     /note="GRAIL 2 excellent exon, frame 0"
     misc_feature    10781. .11209
                     /note="GRAIL 2 excellent exon, frame 2"
     repeat_region   17963. .18120
                     /rpt_family="Alu"
     repeat_region   18993. .19031
                     /note="(ATG)13"
                     /rpt_type=tandem
                     /rpt_unit=ATG
     repeat_region   19242. .19279
                     /note="(AC)19"
                     /rpt_type=tandem
                     /rpt_unit=AC
     misc_feature    complement(21330. .21506)
                     /note="GRAIL 2 excellent exon, frame 1"
     repeat_region   23921. .23956
                     /note="(GT)18"
                     /rpt_type=tandem
                     /rpt_unit=GT
     repeat_region   26847. .26868
                     /note="(T)22"
                     /rpt_type=tandem
                     /rpt_unit=T
     repeat_region   31377. .31420
                     /note="(CT)22"
                     /rpt_type=tandem
                     /rpt_unit=CT
     repeat_region   31442. .31475
                     /note="(AG)17"
                     /rpt_type=tandem
                     /rpt_unit=AG
     repeat_region   31706. .31743
                     /note="(AC)19"
                     /rpt_type=tandem
```

```
                /rpt_unit=AC
repeat_region   31744. .31793
                /note="(AG)25"
                /rpt_type=tandem
                /rpt_unit=AG
repeat_region   32283. .32431
                /rpt_family="Alu"
repeat_region   34229. .34356
                /rpt_family="Alu"
repeat_region   34339. .34359
                /note="(A)21"
                /rpt_type=tandem
                /rpt_unit=A
repeat_region   36102. .36260
                /rpt_family="Alu"
misc_feature    36281. .36383
                /note="GRAIL 2 excellent exon, frame 2"
repeat_region   complement(37129. .37274)
                /rpt_family="Alu"
repeat_region   complement(39179. .39328)
                /rpt_family="Alu"
repeat_region   43238. .43261
                /note="(GTTT)6"
                /rpt_type=tandem
                /rpt_unit=GTTT
misc_feature    43599. .43872
                /note="GRAIL 2 excellent exon, frame 2"
repeat_region   44782. .44823
                /note="(GT)21"
                /rpt_type=tandem
                /rpt_unit=GT
repeat_region   45180. .45207
                /note="(GT)14"
                /rpt_type=tandem
                /rpt_unit=GT
repeat_region   47742. .47765
                /note="(GT)12"
                /rpt_type=tandem
                /rpt_unit=GT
misc_feature    51912. .52056
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    52168. .52252
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    52796. .52901
                /note="GRAIL 2 excellent exon, frame 1"
repeat_region   53966. .53993
                /note="(TTGT)7"
                /rpt_type=tandem
                /rpt_unit=TTGT
misc_feature    56295. .56434
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    56883. .57080
                /note="GRAIL 2 excellent exon, frame 2"
misc_feature    57358. .57513
                /note="GRAIL 2 excellent exon, frame 2"
misc_feature    59918. .59998
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   61088. .61169
                /rpt_family="Alu"
misc_feature    63644. .63822
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    66362. .66598
                /note="GRAIL 2 excellent exon, frame 2"
misc_feature    66987. .67185
                /note="GRAIL 2 excellent exon, frame 0"
misc_feature    67345. .67485
                /note="GRAIL 2 excellent exon, frame 0"
misc_feature    67912. .68019
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   70200. .70225
                /note="(GT)13"
                /rpt_type=tandem
                /rpt_unit=GT
misc_feature    71363. .71431
                /note="GRAIL 2 excellent exon, frame 1"
misc_feature    71776. .71900
                /note="GRAIL 2 excellent exon, frame 0"
misc_feature    complement(73887. .73930)
                /note="GRAIL 2 excellent exon, frame 1"
repeat_region   79838. .79899
                /note="(TC)31"
                /rpt_type=tandem
                /rpt_unit=TC
repeat_region   80689. .80841
                /rpt_family="Alu"
repeat_region   83589. .83612
                /note="(AT)12"
                /rpt_type=tandem
                /rpt_unit=AT
repeat_region   84528. .84579
                /note="(AT)26"
                /rpt_type=tandem
                /rpt_unit=AT
repeat_region   85024. .85065
                /note="(GA)21"
                /rpt_type=tandem
                /rpt_unit=GA
misc_feature    86007. .86052
                /note="GRAIL 2 excellent exon, frame 0"
repeat_region   86454. .86489
                /note="(CTTG)9"
                /rpt_type=tandem
                /rpt_unit=CTTG
repeat_region   complement(86625. .86682)
```

```
  Query Match             41.3%;  Score 31;  DB 12;  Length 159500;
  Best Local Similarity   78.7%;  Pred. No. 38;
  Matches   37;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      10 TTGAAATTTTATTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGG 56
           ||||| |||| ||||| |||||||||| ||  || |||||| |  ||
Db 107870 TTGAATTTTTTTTTTCTTTTTTGGAAAATCCAGTGGCCGCTAAAGG 107916


RESULT   9
AC020886
LOCUS       AC020886   196382 bp    DNA             HTG       16-FEB-2000
DEFINITION  Mus musculus clone RP23-464K4, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC020886
VERSION     AC020886.2  GI:6980213
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 196382)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196382)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 16, 2000 this sequence version replaced gi:6686422.
            * NOTE: This record contains 144 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
```

```
*        1      855: contig of 855 bp in length
*                    gap of unknown length
*      856     1440: contig of 585 bp in length
*                    gap of unknown length
*     1441     1718: contig of 278 bp in length
*                    gap of unknown length
*     1719     2205: contig of 487 bp in length
*                    gap of unknown length
*     2206     2703: contig of 498 bp in length
*                    gap of unknown length
*     2704     3509: contig of 806 bp in length
*                    gap of unknown length
*     3510     4079: contig of 570 bp in length
*                    gap of unknown length
*     4080     4558: contig of 479 bp in length
*                    gap of unknown length
*     4559     4812: contig of 254 bp in length
*                    gap of unknown length
*     4813     5923: contig of 1111 bp in length
*                    gap of unknown length
*     5924     6901: contig of 978 bp in length
*                    gap of unknown length
*     6902     7515: contig of 614 bp in length
*                    gap of unknown length
*     7516     8535: contig of 1020 bp in length
*                    gap of unknown length
*     8536     8750: contig of 215 bp in length
*                    gap of unknown length
*     8751     9373: contig of 623 bp in length
*                    gap of unknown length
*     9374    10192: contig of 819 bp in length
*                    gap of unknown length
*    10193    10880: contig of 688 bp in length
*                    gap of unknown length
*    10881    11050: contig of 170 bp in length
*                    gap of unknown length
*    11051    11269: contig of 219 bp in length
*                    gap of unknown length
*    11270    12430: contig of 1161 bp in length
*                    gap of unknown length
*    12431    13216: contig of 786 bp in length
*                    gap of unknown length
*    13217    13642: contig of 426 bp in length
*                    gap of unknown length
*    13643    14688: contig of 1046 bp in length
*                    gap of unknown length
*    14689    14797: contig of 109 bp in length
*                    gap of unknown length
*    14798    15653: contig of 856 bp in length
*                    gap of unknown length
*    15654    16374: contig of 721 bp in length
*                    gap of unknown length
*    16375    17396: contig of 1022 bp in length
*                    gap of unknown length
*    17397    17504: contig of 108 bp in length
*                    gap of unknown length
*    17505    18320: contig of 816 bp in length
*                    gap of unknown length
*    18321    19338: contig of 1018 bp in length
*                    gap of unknown length
*    19339    20614: contig of 1276 bp in length
*                    gap of unknown length
*    20615    21816: contig of 1202 bp in length
*                    gap of unknown length
*    21817    22751: contig of 935 bp in length
*                    gap of unknown length
*    22752    23821: contig of 1070 bp in length
*                    gap of unknown length
*    23822    24608: contig of 787 bp in length
*                    gap of unknown length
*    24609    24769: contig of 161 bp in length
*                    gap of unknown length
*    24770    25698: contig of 929 bp in length
*                    gap of unknown length
*    25699    26800: contig of 1102 bp in length
*                    gap of unknown length
*    26801    27624: contig of 824 bp in length
*                    gap of unknown length
*    27625    27848: contig of 224 bp in length
*                    gap of unknown length
*    27849    28500: contig of 652 bp in length
*                    gap of unknown length
*    28501    29386: contig of 886 bp in length
*                    gap of unknown length
*    29387    30483: contig of 1097 bp in length
*                    gap of unknown length
*    30484    30732: contig of 249 bp in length
*                    gap of unknown length
*    30733    32258: contig of 1526 bp in length
*                    gap of unknown length
*    32259    32965: contig of 707 bp in length
*                    gap of unknown length
*    32966    33539: contig of 574 bp in length
*                    gap of unknown length
*    33540    34630: contig of 1091 bp in length
*                    gap of unknown length
*    34631    35567: contig of 937 bp in length
*                    gap of unknown length
*    35568    35664: contig of 97 bp in length
*                    gap of unknown length
*    35665    36333: contig of 669 bp in length
*                    gap of unknown length
*    36334    36716: contig of 383 bp in length
*                    gap of unknown length
*    36717    37117: contig of 401 bp in length
*                    gap of unknown length
*    37118    38276: contig of 1159 bp in length
*                    gap of unknown length
*    38277    39282: contig of 1006 bp in length
*                    gap of unknown length
*    39283    39653: contig of 371 bp in length
*                    gap of unknown length
*    39654    40167: contig of 514 bp in length
*                    gap of unknown length
*    40168    40870: contig of 703 bp in length
*                    gap of unknown length
*    40871    41129: contig of 259 bp in length
*                    gap of unknown length
*    41130    41399: contig of 270 bp in length
*                    gap of unknown length
*    41400    41653: contig of 254 bp in length
*                    gap of unknown length
*    41654    42290: contig of 637 bp in length
*                    gap of unknown length
*    42291    43331: contig of 1041 bp in length
*                    gap of unknown length
*    43332    44499: contig of 1168 bp in length
*                    gap of unknown length
*    44500    45695: contig of 1196 bp in length
*                    gap of unknown length
*    45696    46507: contig of 812 bp in length
*                    gap of unknown length
*    46508    47522: contig of 1015 bp in length
*                    gap of unknown length
*    47523    47757: contig of 235 bp in length
*                    gap of unknown length
*    47758    48001: contig of 244 bp in length
*                    gap of unknown length
*    48002    48611: contig of 610 bp in length
*                    gap of unknown length
*    48612    49155: contig of 544 bp in length
*                    gap of unknown length
*    49156    49255: contig of 100 bp in length
*                    gap of unknown length
*    49256    50108: contig of 853 bp in length
*                    gap of unknown length
```

```
            *    50109    51353: contig of 1245 bp in length
            *             gap of unknown length
            *    51354    51530: contig of 177 bp in length
            *             gap of unknown length
            *    51531    51937: contig of 407 bp in length
            *             gap of unknown length
            *    51938    52066: contig of 129 bp in length
            *             gap of unknown length
            *    52067    52220: contig of 154 bp in length
            *             gap of unknown length
            *    52221    52939: contig of 719 bp in length
            *             gap of unknown length
            *    52940    53795: contig of 856 bp in length
            *             gap of unknown length
            *    53796    55050: contig of 1255 bp in length
            *             gap of unknown length
            *    55051    56030: contig of 980 bp in length
            *             gap of unknown length
            *    56031    57128: contig of 1098 bp in length
            *             gap of unknown length
            *    57129    57705: contig of 577 bp in length
            *             gap of unknown length
            *    57706    58825: contig of 1120 bp in length
            *             gap of unknown length

  Query Match             41.3%;  Score 31;  DB 55;  Length 196382;
  Best Local Similarity   78.7%;  Pred. No. 39;
  Matches   37;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      10 TTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGG 56
           ||||| |||| ||||| |||||||||| ||  || |||||| |  ||
Db 194088 TTGAATTTTTTTTTTCTTTTTTGGAAAATCCAGTGGCCGCTAAAGG 194134


RESULT  10
AC006895
LOCUS       AC006895   201746 bp    DNA             HTG       26-FEB-1999
DEFINITION  Caenorhabditis elegans clone Y71G12, *** SEQUENCING IN PROGRESS
            ***, 12 unordered pieces.
ACCESSION   AC006895
VERSION     AC006895.2  GI:4309904
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
            Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 201746)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Caehorhabditis elegans clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201746)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 1, 1999 this sequence version replaced gi:4263453.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2190: contig of 2190 bp in length
            *     2191     2205: gap of unknown length
            *     2206     4263: contig of 2058 bp in length
            *     4264     4278: gap of unknown length
            *     4279     6646: contig of 2368 bp in length
            *     6647     6661: gap of unknown length
            *     6662    10186: contig of 3525 bp in length
            *    10187    10201: gap of unknown length
            *    10202    14198: contig of 3997 bp in length
            *    14199    14213: gap of unknown length
            *    14214    18232: contig of 4019 bp in length
            *    18233    18247: gap of unknown length
            *    18248    23677: contig of 5430 bp in length
            *    23678    23692: gap of unknown length
            *    23693    29225: contig of 5533 bp in length
            *    29226    29240: gap of unknown length
            *    29241    37388: contig of 8148 bp in length
            *    37389    37403: gap of unknown length
            *    37404    69862: contig of 32459 bp in length
            *    69363    69877: gap of unknown length
            *    69378   128770: contig of 58893 bp in length
            *   128771   128785: gap of unknown length
            *   128786   201746: contig of 72961 bp in length.
FEATURES             Location/Qualifiers
     source          1. .201746
                     /organism="Caenorhabditis elegans"
                     /db_xref="taxon:6239"
                     /clone="Y71G12"
BASE COUNT    63828 a  37036 c  36424 g  64293 t    165 others
ORIGIN


  Query Match             40.5%;  Score 30.4;  DB 41;  Length 201746;
  Best Local Similarity   85.0%;  Pred. No. 59;
  Matches   34;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
           | |||||||| |||||| |||||||||||| | ||| |||
Db 185547 CAAAAAATTGTAATTTTTTTTTTTTTTTTTCGGATAAAAA 185586


RESULT  11
AC005341/c
LOCUS       AC005341   161449 bp    DNA             PRI       29-JUL-1999
DEFINITION  Homo sapiens chromosome 17, clone hRPK.146_P_2, complete sequence.
ACCESSION   AC005341
VERSION     AC005341.12  GI:5649369
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161449)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone hRPK.146_P_2
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 161449)
  AUTHORS   Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatin,C.,
            Boutwell.C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E.,
            Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P.,
            FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
            Gensheimer,S., Geraigery,K., Gilmartin,T., Grant,G., Hagos,B.,
            Harris,K., Horton,L., Howland,J.C., Jacotot,L., Kann,L.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B., Peterson,K.,
            Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R.,
            Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,
            Subramanian,A., Torruella-Miller,I., Vassiliev,H., Vo,A.,
            Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-1998) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 161449)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
```

Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (29-JUL-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jul 29, 1999 this sequence version replaced gi:5542046.
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.

FEATURES Location/Qualifiers

source	1. .161449
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="hRPK.146_P_2"
	/clone_lib="RPCI-11 human BAC library"
	/map="17"
	/chromosome="17"
repeat_region	complement(810. .949)
	/rpt_family="MER5B"
repeat_region	1332. .1356
	/rpt_family="AT_rich"
repeat_region	complement(1925. .2026)
	/rpt_family="MER5B"
repeat_region	2810. .2896
	/rpt_family="MIR"
repeat_region	complement(3495. .3674)
	/rpt_family="MIR"
repeat_region	complement(4404. .4514)
	/rpt_family="L2"
repeat_region	4955. .4986
	/rpt_family="(CAT)n"
repeat_region	5500. .5529
	/rpt_family="AT_rich"
repeat_region	complement(5930. .8561)
	/rpt_family="Charlie3"
repeat_region	8589. .8612
	/rpt_family="AT_rich"
repeat_region	8613. .8907
	/rpt_family="AluSg"
repeat_region	8908. .8918
	/rpt_family="AT_rich"
repeat_region	9126. .9167
	/rpt_family="AT_rich"
repeat_region	10187. .10698
	/rpt_family="MER9"
repeat_region	10920. .10980
	/rpt_family="CT-rich"
repeat_region	complement(11178. .11264)
	/rpt_family="MIR"
repeat_region	11585. .11637
	/rpt_family="AT_rich"
repeat_region	11638. .11702
	/rpt_family="(TA)n"
repeat_region	11920. .11947
	/rpt_family="AT_rich"
repeat_region	complement(12020. .12514)
	/rpt_family="L1MD1"
repeat_region	12619. .12652
	/rpt_family="AT_rich"
repeat_region	12761. .12787
	/rpt_family="AT_rich"
repeat_region	12867. .12904
	/rpt_family="AT_rich"
repeat_region	12910. .12995
	/rpt_family="(TA)n"
repeat_region	complement(12996. .13300)
	/rpt_family="AluSx"
repeat_region	13301. .13335
	/rpt_family="(TA)n"
repeat_region	13756. .13824
	/rpt_family="AT_rich"
repeat_region	13879. .13923
	/rpt_family="AT_rich"
repeat_region	14075. .14117
	/rpt_family="(TA)n"
repeat_region	14365. .14398
	/rpt_family="(TTTA)n"
repeat_region	complement(14399. .15428)
	/rpt_family="L1PA2"
repeat_region	15428. .16904
	/rpt_family="L1PA2"
repeat_region	17531. .17552
	/rpt_family="AT_rich"
repeat_region	18835. .19103
	/rpt_family="MLT1K"
repeat_region	19145. .19442
	/rpt_family="AluY"
repeat_region	complement(19501. .19824)
	/rpt_family="L2"
repeat_region	complement(21666. .21766)
	/rpt_family="L2"
repeat_region	22391. .22429
	/rpt_family="AT_rich"
repeat_region	complement(22499. .22852)
	/rpt_family="THE1C"
repeat_region	22855. .23102
	/rpt_family="L1M4c"
repeat_region	24001. .24413
	/rpt_family="L1M4"
repeat_region	complement(25099. .25287)
	/rpt_family="MER53"
repeat_region	25489. .25805
	/rpt_family="AluJb"
repeat_region	25807. .25837
	/rpt_family="(GAAAA)n"
repeat_region	25957. .26580
	/rpt_family="L1M4"
repeat_region	complement(26673. .26939)
	/rpt_family="L2"
repeat_region	27628. .27653
	/rpt_family="(TTA)n"
repeat_region	complement(27655. .27936)
	/rpt_family="AluSx"
repeat_region	complement(28657. .28726)
	/rpt_family="MER5A"
repeat_region	28853. .29041
	/rpt_family="L2"
repeat_region	29215. .29327
	/rpt_family="(TA)n"
repeat_region	complement(29601. .29878)
	/rpt_family="L1"
repeat_region	complement(30090. .30704)
	/rpt_family="MLT2I"
repeat_region	30811. .30840
	/rpt_family="AT_rich"
repeat_region	31465. .31499
	/rpt_family="AT_rich"
repeat_region	complement(31797. .32050)
	/rpt_family="L2"
repeat_region	complement(33538. .33697)
	/rpt_family="MER63"
repeat_region	33837. .34325
	/rpt_family="MER68"
repeat_region	35093. .35135
	/rpt_family="AT_rich"
repeat_region	35426. .35483
	/rpt_family="AT_rich"
repeat_region	37991. .38011
	/rpt_family="AT_rich"

```
     repeat_region   38460. .38507
                     /rpt_family="(TG)n"
     repeat_region   38618. .38915
                     /rpt_family="AluSx"
     repeat_region   complement(40279. .40618)
                     /rpt_family="MER46B"
     repeat_region   complement(40768. .41251)
                     /rpt_family="L1MA8"
     repeat_region   41252. .41284
                     /rpt_family="(TG)n"
     repeat_region   41303. .41388
                     /rpt_family="(TA)n"
     repeat_region   complement(41390. .41557)
                     /rpt_family="L1MA7"
     repeat_region   41558. .41673
                     /rpt_family="(TA)n"

  Query Match             40.3%;  Score 30.2;  DB 40;  Length 161449;
  Best Local Similarity   74.5%;  Pred. No. 66;
  Matches   38;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 CAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATG 54
          ||| |  |||||| ||||||||||||||| ||||  || ||   |  |||||
Db 101906 CAATACCTGAAATCTTATTTTTTTTTTTTCGGAAAGTATAGGTTCTTCCATG 101856


RESULT  12
AC021672/c
LOCUS       AC021672   161986 bp    DNA             HTG       19-JAN-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-88I22 map 11, *** SEQUENCING
            IN PROGRESS ***, 35 unordered pieces.
ACCESSION   AC021672
VERSION     AC021672.1  GI:6716080
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161986)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 11, clone RP11-88I22
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 161986)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1363
                Center clone name: 88_I_22
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 35 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1708: contig of 1708 bp in length
            *                   gap of unknown length
            *     1709     3194: contig of 1486 bp in length
            *                   gap of unknown length
            *     3195     4663: contig of 1469 bp in length
            *                   gap of unknown length
            *     4664     6061: contig of 1398 bp in length
            *                   gap of unknown length
            *     6062     7790: contig of 1729 bp in length
            *                   gap of unknown length
            *     7791     9714: contig of 1924 bp in length
            *                   gap of unknown length
            *     9715    12170: contig of 2456 bp in length
            *                   gap of unknown length
            *    12171    14485: contig of 2315 bp in length
            *                   gap of unknown length
            *    14486    17080: contig of 2595 bp in length
            *                   gap of unknown length
            *    17081    19657: contig of 2577 bp in length
            *                   gap of unknown length
            *    19658    22929: contig of 3272 bp in length
            *                   gap of unknown length
            *    22930    25847: contig of 2918 bp in length
            *                   gap of unknown length
            *    25848    27994: contig of 2147 bp in length
            *                   gap of unknown length
            *    27995    31034: contig of 3040 bp in length
            *                   gap of unknown length
            *    31035    33767: contig of 2733 bp in length
            *                   gap of unknown length
            *    33768    36270: contig of 2503 bp in length
            *                   gap of unknown length
            *    36271    39201: contig of 2931 bp in length
            *                   gap of unknown length
            *    39202    42313: contig of 3112 bp in length
            *                   gap of unknown length
            *    42314    44740: contig of 2427 bp in length
            *                   gap of unknown length
            *    44741    48372: contig of 3632 bp in length
            *                   gap of unknown length
            *    48373    54434: contig of 6062 bp in length
            *                   gap of unknown length
            *    54435    59506: contig of 5072 bp in length
            *                   gap of unknown length
            *    59507    63411: contig of 3905 bp in length
            *                   gap of unknown length
            *    63412    66980: contig of 3569 bp in length
            *                   gap of unknown length
            *    66981    70955: contig of 3975 bp in length
            *                   gap of unknown length
            *    70956    75938: contig of 4983 bp in length
            *                   gap of unknown length
            *    75939    81365: contig of 5427 bp in length
            *                   gap of unknown length
            *    81366    87661: contig of 6296 bp in length
            *                   gap of unknown length
            *    87662    95607: contig of 7946 bp in length
            *                   gap of unknown length
            *    95608   104091: contig of 8484 bp in length
            *                   gap of unknown length
            *   104092   114039: contig of 9948 bp in length
            *                   gap of unknown length
            *   114040   123090: contig of 9051 bp in length
```

```
            *                    gap of unknown length
            *   123091   132651: contig of 9561 bp in length
            *                    gap of unknown length
            *   132652   144718: contig of 12067 bp in length
            *                    gap of unknown length
            *   144719   161986: contig of 17268 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161986
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11"
                     /clone="RP11-88I22"
                     /clone_lib="RPCI-11 Human Male BAC"
BASE COUNT    45594 a  36630 c  35856 g  43893 t     13 others
ORIGIN


  Query Match           40.3%;  Score 30.2;  DB 45;  Length 161986;
  Best Local Similarity  81.4%;  Pred. No. 66;
  Matches   35;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAG 43
           |||||||| ||||| |||| ||||||||||||    | |||||
Db   34152 GGCCAAAAGTTGAAGTTTTTTTTTTTTTTTTTTTTTTTTAAAG 34110


RESULT  13
AC009526
LOCUS       AC009526   108061 bp    DNA             PLN       11-DEC-1999
DEFINITION  Arabidopsis thaliana chromosome I BAC F2J6 genomic sequence,
            complete sequence.
ACCESSION   AC009526
VERSION     AC009526.3  GI:6560894
KEYWORDS    HTG.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
            eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
            Arabidopsis.
REFERENCE   1  (bases 1 to 108061)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 108061)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-1999) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 108061)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-1999) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     On Dec 11, 1999 this sequence version replaced gi:5881522.
            Bases 1-4149 of IGF clone F2J6 overlap with bases 91042-95190 of
            TAMU clone T10P12 (gb|AC007203) and an unknown number of bases
            overlap at the end of F2J6 overlap with IGF clone F28H19
            (gb|AC006423).
FEATURES             Location/Qualifiers
     source          1. .108061
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="I"
                     /clone="F2J6"
BASE COUNT    35398 a  19039 c  19111 g  34509 t      4 others
ORIGIN


  Query Match           39.7%;  Score 29.8;  DB 8;  Length 108061;
  Best Local Similarity  82.9%;  Pred. No. 82;
  Matches   34;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 GCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
           || ||||||  |||||  |||||||||| |||||||| ||
Db   25278 GCTAAAAATGAAAATTCAATTTTTTTTTGTTGGAATATTAA 25318


RESULT  14
HS1093G12/c
LOCUS       HS1093G12  141655 bp    DNA             HTG       04-FEB-2000
DEFINITION  Homo sapiens chromosome 20 clone RP5-1093G12, *** SEQUENCING IN
            PROGRESS ***, in unordered pieces.
ACCESSION   AL121751
VERSION     AL121751.7  GI:6630762
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 141655)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 22, 1999 this sequence version replaced gi:6562044.
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence.  Work on the sequence is in
            progress and the release of this data is based on the understanding
            that the sequence may change as work continues.  The sequence may
            be contaminated with foreign sequence from E.coli, yeast, vector,
            phage etc. Order of segments is not known; 800 n's separate
            segments. Contig_ID: 00409  Length: 141655bp.
            * NOTE: This is a 'working draft' sequence.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
FEATURES             Location/Qualifiers
     source          1. .141655
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP5-1093G12"
                     /clone_lib="RPCI-5"
BASE COUNT    39105 a  32980 c  31848 g  37722 t
ORIGIN


  Query Match           39.7%;  Score 29.8;  DB 32;  Length 141655;
  Best Local Similarity  70.2%;  Pred. No. 85;
  Matches   40;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       6 AAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGC 62
           |||| || ||| ||||||||||| ||| || | |||   |||    || ||| |||
Db 122840 AAAAATGGAATATTATTTTTTTTCTTTTAAAAAAAAATATGCCAAGCTGCGCCAGGC 122784


RESULT  15
HSJ324O17
LOCUS       HSJ324O17  218305 bp    DNA             HTG       17-FEB-2000
DEFINITION  Homo sapiens chromosome 20 clone RP3-324O17, *** SEQUENCING IN
```

```
            PROGRESS ***, 24 unordered pieces.
ACCESSION   AL110115
VERSION     AL110115.16  GI:7007308
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 218305)
  AUTHORS   Dunn,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 19, 2000 this sequence version replaced gi:6983401.
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence.  Work on the sequence is in
            progress and the release of this data is based on the understanding
            that the sequence may change as work continues.  The sequence may
            be contaminated with foreign sequence from E.coli, yeast, vector,
            phage etc. Order of segments is not known; 800 n's separate
            segments. Contig_ID: 00211  Length: 65856bp
            Contig_ID: 00229  Length: 22806bp
            Contig_ID: 01346  Length: 2445bp
            Contig_ID: 01393  Length: 1021bp
            Contig_ID: 01651  Length: 29038bp
            Contig_ID: 01653  Length: 5553bp
            Contig_ID: 02437  Length: 7099bp
            Contig_ID: 02495  Length: 5023bp
            Contig_ID: 02585  Length: 1081bp
            Contig_ID: 02718  Length: 1504bp
            Contig_ID: 02932  Length: 1607bp
            Contig_ID: 02962  Length: 4240bp
            Contig_ID: 02978  Length: 4486bp
            Contig_ID: 03061  Length: 1109bp
            Contig_ID: 03110  Length: 2159bp
            Contig_ID: 03114  Length: 1405bp
            Contig_ID: 03137  Length: 1246bp
            Contig_ID: 03149  Length: 1994bp
            Contig_ID: 03157  Length: 3860bp
            Contig_ID: 03199  Length: 5167bp
            Contig_ID: 03226  Length: 1158bp
            Contig_ID: 03235  Length: 1225bp
            Contig_ID: 03242  Length: 2725bp
            Contig_ID: 03251  Length: 26098bp.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 24 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    65856: contig of 65856 bp in length
            *   65857 66656: gap of      800 bp
            *   66657    89462: contig of 22806 bp in length
            *   89463 90262: gap of      800 bp
            *   90263    92707: contig of 2445 bp in length
            *   92708 93507: gap of      800 bp
            *   93508    94528: contig of 1021 bp in length
            *   94529 95328: gap of      800 bp
            *   95329   124366: contig of 29038 bp in length
            *  124367 125166: gap of      800 bp
            *  125167   130719: contig of 5553 bp in length
            *  130720 131519: gap of      800 bp
            *  131520   138618: contig of 7099 bp in length
            *  138619 139418: gap of      800 bp
            *  139419   144441: contig of 5023 bp in length
            *  144442 145241: gap of      800 bp
            *  145242   146322: contig of 1081 bp in length
            *  146323 147122: gap of      800 bp
            *  147123   148626: contig of 1504 bp in length
            *  148627 149426: gap of      800 bp
            *  149427   151033: contig of 1607 bp in length
            *  151034 151833: gap of      800 bp
            *  151834   156073: contig of 4240 bp in length
            *  156074 156873: gap of      800 bp
            *  156874   161359: contig of 4486 bp in length
            *  161360 162159: gap of      800 bp
            *  162160   163268: contig of 1109 bp in length
            *  163269 164068: gap of      800 bp
            *  164069   166227: contig of 2159 bp in length
            *  166228 167027: gap of      800 bp
            *  167028   168432: contig of 1405 bp in length
            *  168433 169232: gap of      800 bp
            *  169233   170478: contig of 1246 bp in length
            *  170479 171278: gap of      800 bp
            *  171279   173272: contig of 1994 bp in length
            *  173273 174072: gap of      800 bp
            *  174073   177932: contig of 3860 bp in length
            *  177933 178732: gap of      800 bp
            *  178733   183899: contig of 5167 bp in length
            *  183900 184699: gap of      800 bp
            *  184700   185857: contig of 1158 bp in length
            *  185858 186657: gap of      800 bp
            *  186658   187882: contig of 1225 bp in length
            *  187883 188682: gap of      800 bp
            *  188683   191407: contig of 2725 bp in length
            *  191408 192207: gap of      800 bp
            *  192208   218305: contig of 26098 bp in length.
FEATURES             Location/Qualifiers
     source          1. .218305
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP3-324O17"
                     /clone_lib="RPCI-3"
BASE COUNT    51570 a  47647 c  48369 g  52311 t  18408 others
ORIGIN


  Query Match             39.7%;  Score 29.8;  DB 33;  Length 218305;
  Best Local Similarity   70.2%;  Pred. No. 89;
  Matches   40;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      6 AAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGC 62
          |||| || ||| ||||||||||| |||  || | |||   |||    || ||| |||
Db 202922 AAAAATGGAATATTATTTTTTTTCTTTTAAAAAAAAATATGCCAAGCTGCGCCAGGC 202978


Search completed: May 29, 2000, 21:35:19
Job time: 36851 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:06 ; Search time 1446.77 Seconds
(without alignments)
12.970 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GGCCAAAAATTGAAATTTTA.....GCCCGGCCGCCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		40	53.3	84	1	T78340	Chimeric virus con
c 2		40	53.3	86	1	T78341	Chimeric virus con
c 3		39.6	52.8	6474	1	Q40300	Sequence of plasmi
c 4		38.8	51.7	5472	1	Q41005	Sequence of plasmi
c 5		38.4	51.2	6811	1	Q41004	Sequence of plasmi
6		38	50.7	42	1	Q67969	Sequence of modifi
7		38	50.7	51	1	Q40295	Sequence of oligo
c 8		38	50.7	53	1	Q40296	Sequence of oligo
c 9		38	50.7	3878	1	Q40299	Sequence of plasmi
10		27	36.0	46	1	Q41030	pFP-Zsart construc
11		27	36.0	62	1	Q41025	P11 promoter mutan
12		27	36.0	88	1	Q40248	Sequence of plasmi
c 13		27	36.0	92	1	Q40249	Sequence of plasmi
c 14		26.6	35.5	97	1	Q40278	Sequence of synthe
15		26.6	35.5	134	1	Q40250	Sequence of a DNA
16		26.6	35.5	4145	1	Q40279	Sequence of clone
c 17		26.6	35.5	25464	1	V57274	Human flavin-conta
18		26	34.7	2002	1	V36503	Human D2-2 gene as
19		25.8	34.4	90	1	V37789	Analytical solid p
20		25.8	34.4	4093	1	V58357	Coding sequence fo
c 21		25.6	34.1	42	1	Q41029	pFP-Zsart construc
c 22		25.4	33.9	126	1	V89886	EST clone CT317. N
c 23		25.4	33.9	3471	1	T71761	PSKH-1 cDNA. New i
24		25.2	33.6	55	1	Q41006	Sequence of modifi
c 25		25.2	33.6	232	1	Q77192	Human genome fragm
c 26		25.2	33.6	3095	1	T31290	Mouse poly-immunog
c 27		25	33.3	1923	1	X51745	DNA encoding a hum
28		25	33.3	2453	1	V52449	Streptococcus pneu
c 29		25	33.3	3256	1	V29063	BRCA1 modulator pr
30		25	33.3	13425	1	V52284	Streptococcus pneu
31		25	33.3	16891	1	X37084	MEFV gene sequence
c 32		24.8	33.1	2436	1	V20438	Human c-N-ras onco
33		24.6	32.8	238	1	T26771	Human gene signatu
c 34		24.6	32.8	711	1	X00670	Human secreted pro
35		24.6	32.8	2058	1	V07560	Neocallimastix pat
c 36		24.6	32.8	3288	1	T05055	Human transforming
c 37		24.6	32.8	5040	1	N80916	Sequence encoding
38		24.4	32.5	91	1	Q41021	P11 promoter wildt
c 39		24.4	32.5	768	1	X39988	Prostate cancer as
c 40		24.4	32.5	812	1	X39990	Prostate cancer as
41		24.2	32.3	368	1	V87415	EST clone BR733. N
42		24.2	32.3	1296	1	V34269	Human secreted pro
43		24.2	32.3	5760	1	N50530	Sequence encoding
44		24.2	32.3	110000	1	V21209_13	Continuation (14 o
c 45		24	32.0	981	1	V60576	ERAB protein codin

ALIGNMENTS

RESULT 1
T78340
ID T78340 standard; DNA; 84 BP.
AC T78340;
DT 13-OCT-1997 (first entry)
DE Chimeric virus construction oligonucleotide oselP1.
KW Recombinant eukaryotic cytoplasmic DNA virus; vaccine production;
KW double selection marker; attenuated live pox virus; HIV; HBV;
KW surface antigen gene; encephalitis; tick-borne; ss.
OS Synthetic.
PN EP-753581-A1.
PD 15-JAN-1997.
PF 10-JUL-1995; 110727.
PR 10-JUL-1995; EP-110727.
PA (IMMO) IMMUNO AG.
PI Antoine G, Dorner F, Eibl J, Falkner F, Scheiflinger F;
DR WPI; 97-079382/08.
PT Prodn. of recombinant eukaryotic cytoplasmic DNA viruses for vaccine
PT prodn. - using novel construct contg. foreign DNA and double
PT selection marker
PS Example 9; Page 14; 68pp; English.
CC A method has been produced for the production of recombinant eukaryotic
CC cytoplasmic DNA viruses for vaccine production. The method involves
CC inserting a construct DNA molecule into the genome of a eukaryotic
CC cytoplasmic DNA virus, transfecting cells with the virus, and selecting
CC a recombinant virus that does not contain the double selection marker
CC cassette from the construct DNA. The DNA molecule preferably contains
CC one or more foreign DNA segments of a gene encoding an antigen of a
CC pathogen (especially HIV env, HIV gag, HIV gagpol, HIV nef, HBV
CC preS1-S2-S surface antigen gene, HBV S1-S2-S surface antigen gene, HBV
CC S surface antigen gene, tick-borne encephalitis preM gene or tick-borne
CC encephalitis E gene), and has a double selection marker cassette which
CC is flanked by at least two direct repeat DNA sequences, where at least
CC one foreign DNA segment does not occur within the DNA sequence bounded
CC by the direct repeat sequences and where the double selection marker
CC cassette comprises a dominant selection marker gene and a colour
CC selection marker gene (preferably the colour selection marker is the
CC Escherichia coli lacZ gene and the dominant selection marker is the
CC E.coli hph or gpt gene). The present sequence represents the oselP1
CC oligonucleotide used in the construction of chimeric MVA vaccinia and
CC fowlpox viruses expressing the tick-borne encephalitis virus preM and E
CC genes. The DNA constructs and method are used for the production of
CC vaccines, especially containing an attenuated live recombinant pox
CC virus.
SQ Sequence 84 BP; 28 A; 8 C; 10 G; 38 T;

Query Match 53.3%; Score 40; DB 1; Length 84;
Best Local Similarity 82.1%; Pred. No. 0.0012;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
          | ||||||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db      9 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 64
```

```
RESULT   2
T78341/c
ID   T78341 standard; DNA; 86 BP.
AC   T78341;
DT   13-OCT-1997  (first entry)
DE   Chimeric virus construction oligonucleotide oselP2.
KW   Recombinant eukaryotic cytoplasmic DNA virus; vaccine production;
KW   double selection marker; attenuated live pox virus; HIV; HBV;
KW   surface antigen gene; encephalitis; tick-borne; ss.
OS   Synthetic.
PN   EP-753581-A1.
PD   15-JAN-1997.
PF   10-JUL-1995; 110727.
PR   10-JUL-1995; EP-110727.
PA   (IMMO ) IMMUNO AG.
PI   Antoine G,  Dorner F,  Eibl J,  Falkner F,  Scheiflinger F;
DR   WPI; 97-079382/08.
PT   Prodn. of recombinant eukaryotic cytoplasmic DNA viruses for vaccine
PT   prodn. - using novel construct contg. foreign DNA and double
PT   selection marker
PS   Example 9; Page 14; 68pp; English.
CC   A method has been produced for the production of recombinant eukaryotic
CC   cytoplasmic DNA viruses for vaccine production. The method involves
CC   inserting a construct DNA molecule into the genome of a eukaryotic
CC   cytoplasmic DNA virus, transfecting cells with the virus, and selecting
CC   a recombinant virus that does not contain the double selection marker
CC   cassette from the construct DNA. The DNA molecule preferably contains
CC   one or more foreign DNA segments of a gene encoding an antigen of a
CC   pathogen (especially HIV env, HIV gag, HIV gagpol, HIV nef, HBV
CC   preS1-S2-S surface antigen gene, HBV S1-S2-S surface antigen gene, HBV
CC   S surface antigen gene, tick-borne encephalitis preM gene or tick-borne
CC   encephalitis E gene), and has a double selection marker cassette which
CC   is flanked by at least two direct repeat DNA sequences, where at least
CC   one foreign DNA segment does not occur within the DNA sequence bounded
CC   by the direct repeat sequences and where the double selection marker
CC   cassette comprises a dominant selection marker gene and a colour
CC   selection marker gene (preferably the colour selection marker is the
CC   Escherichia coli lacZ gene and the dominant selection marker is the
CC   E.coli hph or gpt gene). The present sequence represents the oselP2
CC   oligonucleotide used in the construction of chimeric MVA vaccinia and
CC   fowlpox viruses expressing the tick-borne encephalitis virus preM and E
CC   genes. The DNA constructs and method are used for the production of
CC   vaccines, especially containing an attenuated live recombinant pox
CC   virus.
SQ   Sequence   86 BP;    38 A;    11 C;    9 G;    28 T;


  Query Match             53.3%;  Score 40;  DB 1;  Length 86;
  Best Local Similarity   82.1%;  Pred. No. 0.0012;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
          | ||||||||||||||||||||||||||||||||||||||   | | |   ||||
Db     80 CTAAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 25


RESULT   3
Q40300/c
ID   Q40300 standard; DNA; 6474 BP.
AC   Q40300;
DT   02-AUG-1993  (first entry)
DE   Sequence of plasmid pselP-gp160MN.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .55
FT                   /*tag=  a
FT                   /label= pTZ19R
FT   misc_feature    complement (56. .108)
FT                   /*tag=  b
FT                   /label= Linker I in rc orientation
FT   cds             complement (110. .860)
FT                   /*tag=  c
FT                   /label= E. coli gpt
FT   promoter        complement (861. .1245)
FT                   /*tag=  d
FT                   /label= Vaccinia virus p7.5
FT                   /note ="starting with the p7.5 internal NdeI site
FT                   at posn. 1241"
FT   misc_feature    1339. .1344
FT                   /*tag=  e
FT                   /label= HpaI site
FT   misc_feature    complement (1259. .3916)
FT                   /*tag=  f
FT                   /label= HIV-MN env gene
FT   promoter        complement (3917. .3970)
FT                   /*tag=  g
FT                   /label= vaccinia virus synthetic early/late
FT                   promoter
FT   misc_feature    complement (3971. .4015)
FT                   /*tag=  h
FT                   /label= linker II
FT   misc_feature    complement (4016. .6474)
FT                   /*tag=  i
FT                   /label= pTZ19R
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Example; Pages 169-172; 206pp; English.
CC   The synthetic early/late promoter selP was used to express the
CC   gp160-gene of the HIV-1 MN strain.  For the construction of pL2
CC   the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC   by an XbaI-ClaI adaptor fragment consisting of the annealed
CC   oligonucleotide o-542 and o-544.  The intermediate plasmid
CC   resulting from this cloning step was called pL1.  The 0.84kb
CC   AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC   fragment consisting of the annealed oligonucleotide o-541 and
CC   o-543.  The resulting plasmid was called pL2.  The XbaI-SphI
CC   fragment was treated with Klenow-polymerase and inserted between
CC   the PvuIi sites of the plasmid pTZ19R.  The resulting plasmid was
CC   called pTZ-L2.  The 0.6kb ClaI-NcoI fragment (the T7-promoter-EMC-
CC   sequence) was replaced with a synthetic promoter fragment
CC   consisting of the annealed oligonucleotide o-selPI and o-selPI.
CC   The 239bp SaII-NdeI fragment of the resulting intermediate plasmid
CC   was substituted by the SaII-NedI adaptor consisting of the annealed
CC   oligonucleotides o-830 and o-857.  The resulting plasmid was called
CC   pselP-gpt-L2.  The 3.1kb env gene contiaining the EcoRI-PvuII
CC   fragment of pMNenvI was inserted into the EcoRI and StuI cut plasmid
CC   pselP-gpt-L2 resulting in the intermediate plasmid pselP-gp160.1.
CC   The 0.8kb NcoI-NsiI fragment of pselP-gp160 was substituted by a
CC   PCR-generated 0.31kb NcoI-NsiI fragment resulting in the final
CC   plasmid pselP-gp160MN.  The primers used for the PCR reaction were
CC   o-NcoI and o-NsiI.
SQ   Sequence   6474 BP;    1648 A;    1532 C;    1410 G;    1880 T;


  Query Match             52.8%;  Score 39.6;  DB 1;  Length 6474;
  Best Local Similarity   83.3%;  Pred. No. 0.0038;
  Matches   45;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
          |||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db   3964 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 3911


RESULT   4
Q41005/c
ID   Q41005 standard; DNA; 5472 BP.
```

```
AC   Q41005;
DT   02-AUG-1993  (first entry)
DE   Sequence of plasmid pN2gpta-FIX.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .2217
FT                   /*tag= a
FT                   /label= bluescript II SK- sequence
FT   misc_feature    2218. .2225
FT                   /*tag= b
FT                   /label= FIX sequence
FT   cds             complement (2226. .3659)
FT                   /*tag= c
FT                   /label= Prot S
FT   promoter        complement (3660. .3713)
FT                   /*tag= d
FT                   /label= vaccinia virus synthetic early/late
FT                   promoter
FT                   /note= "flanked by a NocI site and a fused Bg1II/
FT                   BamHI site. The NcoI site harbors the start
FT                   codon TAC"
FT   promoter        3714. .4214
FT                   /*tag= e
FT                   /label= vaccinia virus p7.5
FT   cds             4215. .4848
FT                   /*tag= f
FT                   /label= E. coli gpt
FT                   /note= "the ORF starts at posn. 4215 and ends at
FT                   posn. 4671"
FT   misc_feature    4849. .4856
FT                   /*tag= g
FT                   /label= NotI site 2
FT   misc_feature    4857. .5532
FT                   /*tag= h
FT                   /label= bluescript II SK- sequences
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Claim 99; Pages 181-184; 206pp; English.
CC   The example describes the construction of novel chimeric vaccinia
CC   viruses encoding human factor IX and expresssion of recombinant factor
CC   IX.  A double gene cassette consisting of the complementary DNA for the
CC   human blood factor IX and the gpt gene, each controlled by a vaccinia
CC   promoter, was cloned into the unique NotI site of the vaccinia virus WR
CC   genome and packaged in fowlpox helper virus-infected mammalian cells.
CC   Human factor IX was expressed in several cell types.
SQ   Sequence   5472 BP;    1475 A;    1312 C;    1241 G;    1444 T;


  Query Match             51.7%;  Score 38.8;  DB 1;  Length 5472;
  Best Local Similarity   79.3%;  Pred. No. 0.0062;
  Matches   46;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
          | ||||||||||||||||||||||||||||||||||||   | |  |   ||  ||
Db   3649 CTAAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGAGCG 3592


RESULT   5
Q41004/c
ID   Q41004 standard; DNA; 6811 BP.
AC   Q41004;
DT   02-AUG-1993  (first entry)
DE   Sequence of plasmid pN2-gptaProtS.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .2217
FT                   /*tag= a
FT                   /label= bluescript II SK- sequence
FT   misc_feature    2218. .2225
FT                   /*tag= b
FT                   /label= NotI site 1
FT   cds             complement (2226. .4938)
FT                   /*tag= c
FT                   /label= Prot S
FT   promoter        complement (4939. .4992)
FT                   /*tag= d
FT                   /label= vaccinia virus synthetic early/late
FT                   promoter
FT                   /note= "flanked by a NocI site and a fused Bg1II/
FT                   BamHI site. The NcoI site harbors the start
FT                   codon TAC"
FT   promoter        4993. .5493
FT                   /*tag= e
FT                   /label= vaccinia virus p7.5
FT   cds             5494. .6127
FT                   /*tag= f
FT                   /label= E. coli gpt
FT                   /note= "the ORF starts at posn. 5494 and ends at
FT                   posn. 5950"
FT   misc_feature    6228. .6235
FT                   /*tag= g
FT                   /label= NotI site 2
FT   misc_feature    6236. .6811
FT                   /*tag= h
FT                   /label= bluescript II SK- sequences
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Example; Pages 173-176; 206pp; English.
CC   Single-stranded DNA prepared from the plasmid pBluescript-ProtS,
CC   was used to mutagenize the region around the translational start
CC   codon of the protein S coding region.  The signal peptide is
CC   mutated, with the second amino acid changed from Arg to Ala (see
CC   Q41009 and Q41010).  The protein S cDNA was subsequently excised as
CC   an NcoI-NotI fragment and inserted into the vaccinia insertion
CC   plasmid pTKgpt-selP a plasmid providing a strong synthetic vaccinia
CC   promoter.  The promoter-protein S gene cassette was then excised
CC   and inserted into the plasmid pN2-gpta resulting in pN2-gptaProtS.
SQ   Sequence   6811 BP;    1924 A;    1565 C;    1491 G;    1831 T;


  Query Match             51.2%;  Score 38.4;  DB 1;  Length 6811;
  Best Local Similarity   97.5%;  Pred. No. 0.0085;
  Matches   39;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
          | |||||||||||||||||||||||||||||||||||||
Db   4988 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 4949


RESULT   6
Q67969
ID   Q67969 standard; DNA; 42 BP.
AC   Q67969;
DT   02-JAN-1995  (first entry)
DE   Sequence of modified vaccinia promoter p7.5.
KW   Recombinant replicable vaccinia virus; hepatitis; promoter;
```

KW prevention; therapy; epitope; hepatitis B virus; ss.
OS Synthetic.
PN WO9412617-A.
PD 09-JUN-1994.
PF 24-NOV-1993; U11474.
PR 25-NOV-1992; US-982211.
PA (ITBI-) INT BIOTECHNOLOGY LAB INC.
PI Bernstine EG, Lewis T, Okeefe RW, Souw PTS;
DR WPI; 94-200247/24.
PT Prevention and treatment of hepatitis - using recombinant
PT replicable vaccinia viruses contg. hepatitis B virus surface and
PT core antigen nucleotide sequences
PS Disclosure; Fig 6; 252pp; English.
CC HBV core antigen (Ag) encoding sequences were subcloned and
CC engineered so as to be transcriptionally controlled by a
CC vaccinia or vaccinia-like promoter. A strong vaccinia promoter
CC p7.5 (cochran et al., 1985, J. Virol. 54:30-37), having the
CC sequence in Q67967, can be one of several vaccinia or vaccinia-
CC like promoter used to direct expression of the HBV sequences.
CC the p7.5 promoter is so named because it normally directs
CC expression of vaccinia polypeptide of 7.5 kD. Another vaccinia
CC promoter, p11, in nature directs expression of vaccinia structural
CC protein of 11 kD. Unlike p7.5, p11 is active only late in viral
CC replication. Its sequence is given in Q67968. The modified p7.5
CC promoter (Q67969) is a strong synthetic promoter that is active
CC in both early and late viral replication. Its sequence is partly
CC based on the p7.5 promoter.
SQ Sequence 42 BP; 17 A; 0 C; 4 G; 21 T;

Query Match 50.7%; Score 38; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 42
||||||||||||||||||||||||||||||||||||||
Db 1 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 38

RESULT 7
Q40295
ID Q40295 standard; DNA; 51 BP.
AC Q40295;
DT 02-AUG-1993 (first entry)
DE Sequence of oligo o-selPI of a synthetic promoter fragment.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
PN AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO) IMMUNO AG.
PI Dorner F, Falkner FG, Pfleiderer M, Scheiflinger F;
DR WPI; 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example; Page 158; 206pp; English.
CC The synthetic early/late promoter selP was used to express the
CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC by an XbaI-ClaI adaptor fragment consisting of the annealed
CC oligonucleotide o-542 and o-544. The intermediate plasmid
CC resulting from this cloning step was called pL1. The 0.84kb
CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC fragment consisting of the annealed oligonucleotide o-541 and
CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC fragment was treated with Klenow-polymerase and inserted between
CC the PvuIi sites of the plasmid pTZ19R. The resulting plasmid was
CC called pTZ-L2. The 0.6kb ClaI-NcoI fragment (the
CC T7-promoter-EMC-sequence) was replaced with a synthetic promoter
CC fragment consisting of the annealed oligonucleotide o-selPI and
CC o-selPII. The 239bp SaII-NdeI fragment of the resulting
CC intermediate plasmid was substituted by the SalI-NdeI adaptor
CC consisting of the annealed oligonucleotides o-830 and o-857. The
CC resulting plasmid was called pselP-gpt-L2.
SQ Sequence 51 BP; 18 A; 4 C; 6 G; 23 T;

Query Match 50.7%; Score 38; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 42
||||||||||||||||||||||||||||||||||||||
Db 5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 42

RESULT 8
Q40296/c
ID Q40296 standard; DNA; 53 BP.
AC Q40296;
DT 02-AUG-1993 (first entry)
DE Sequence of oligo o-selPII of a synthetic promoter fragment.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
PN AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO) IMMUNO AG.
PI Dorner F, Falkner FG, Pfleiderer M, Scheiflinger F;
DR WPI; 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example; Page 158; 206pp; English.
CC The synthetic early/late promoter selP was used to express the
CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC by an XbaI-ClaI adaptor fragment consisting of the annealed
CC oligonucleotide o-542 and o-544. The intermediate plasmid
CC resulting from this cloning step was called pL1. The 0.84kb
CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC fragment consisting of the annealed oligonucleotide o-541 and
CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC fragment was treated with Klenow-polymerase and inserted between
CC the PvuIi sites of the plasmid pTZ19R. The resulting plasmid was
CC called pTZ-L2. The 0.6kb ClaI-NcoI fragment (the
CC T7-promoter-EMC-sequence) was replaced with a synthetic promoter
CC fragment consisting of the annealed oligonucleotide o-selPI and
CC o-selPII. The 239bp SaII-NdeI fragment of the resulting
CC intermediate plasmid was substituted by the SalI-NdeI adaptor
CC consisting of the annealed oligonucleotides o-830 and o-857. The
CC resulting plasmid was called pselP-gpt-L2.
SQ Sequence 53 BP; 24 A; 6 C; 4 G; 19 T;

Query Match 50.7%; Score 38; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 42
||||||||||||||||||||||||||||||||||||||
Db 51 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAA 14

RESULT 9
Q40299/c
ID Q40299 standard; DNA; 3878 BP.
AC Q40299;
DT 02-AUG-1993 (first entry)
DE Sequence of plasmid pselP-gpt-L2.
KW Plasmid; cloning; restriction site; ss.

```
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .55
FT                   /*tag= a
FT                   /label= pTZ19R
FT   misc_feature    complement (56. .108)
FT                   /*tag= b
FT                   /label= Linker I in rc orientation
FT   cds             complement (110. .860)
FT                   /*tag= c
FT                   /label= E. coli gpt
FT   promoter        complement (861. .1245)
FT                   /*tag= d
FT                   /label= Vaccinia virus p7.5
FT                   /note ="starting with the p7.5 internal NdeI site
FT                   at posn. 1241"
FT   misc_feature    1339. .1344
FT                   /*tag= e
FT                   /label= HpaI site
FT   misc_feature    complement (1259. .1322)
FT                   /*tag= g
FT                   /label= multiple cloning site
FT   promoter        complement (1323. .1374)
FT                   /*tag= i
FT                   /label= vaccinia virus synthetic early/late
FT                   promoter
FT   misc_feature    complement (1375. .1414)
FT                   /*tag= j
FT                   /label= linker II
FT   misc_feature    complement (1415. .3878)
FT                   /*tag= k
FT                   /label= pTZ19R
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F, Falkner FG, Pfleiderer M, Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Example; Pages 165-167; 206pp; English.
CC   The synthetic early/late promoter selP was used to express the
CC   gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC   the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC   by an XbaI-ClaI adaptor fragment consisting of the annealed
CC   oligonucleotide o-542 and o-544. The intermediate plasmid
CC   resulting from this cloning step was called pL1. The 0.84kb
CC   AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC   fragment consisting of the annealed oligonucleotide o-541 and
CC   o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC   fragment was treated with Klenow-polymerase and inserted between
CC   the PvuIi sites of the plasmid pTZ19R. The resulting plasmid was
CC   called pTZ-L2. The 0.6kb ClaI-NcoI fragment (the T7-promoter-EMC-
CC   sequence) was replaced with a synthetic promoter fragment
CC   consisting of the annealed oligonucleotide o-selPI and o-selPI.
CC   The 239bp SaII-NdeI fragment of the resulting intermediate plasmid
CC   was substituted by the SaII-NedI adaptor consisting of the annealed
CC   oligonucleotides o-830 and o-857. The resulting plasmid was called
CC   pselP-gpt-L2.
SQ   Sequence   3878 BP;    1012 A;    927 C;    974 G;    965 T;


  Query Match            50.7%;  Score 38;  DB 1;  Length 3878;
  Best Local Similarity  100.0%;  Pred. No. 0.0099;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
          |||||||||||||||||||||||||||||||||||||
Db   1368 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 1331
```

```
RESULT  10
Q41030
ID   Q41030 standard; DNA; 46 BP.
AC   Q41030;
DT   29-JUL-1993  (first entry)
DE   pFP-Zsart construction oligomer #2.
KW   Fowl pox virus; FPV; P2; promoter; vaccina virus; VV; mutation; apoAI;
KW   multiple cloning site; plasmid; thymidine kinase; tk; Marck's disease;
KW   intergenic region; post-translational modification; factor II; apoAII;
KW   factor V; factor VII; factor VIII; factor IX; tick-borne encephalitis;
KW   hepatitis E; factor X; factor XI; factor XII; factor XIII; protein C;
KW   protein S; von Willebrand factor; hepatitis B; hepatitis C; malaria;
KW   plasminogen; apolipo protein; viral antigen; infectious bronchitis;
KW   TBE; tetanus; HIV; HSV; pertussis; Newcastle disease; vaccine; ss.
OS   Synthetic.
PN   AU9219580-A.
PD   04-MAR-1993.
PF   10-JUL-1992; 019580.
PR   26-AUG-1991; EP-114300.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F, Falkner FG, Scheiflinger F;
DR   WPI; 93-126452/16.
PT   Prepn. of recombinant fowl pox virus for e.g. expression of
PT   proteins - by inserting foreign DNA into enlarged inter-genic
PT   region so FPV TK-gene remains intact and codes for entire
PT   thymidine kinase
PS   Disclosure; Page 39; 92pp; English.
CC   The sequences given in Q41027-45 are oligomers which were used in the
CC   construction of the insertion plasmids of the invention. These
CC   plasmids comprised the fowl pox virus (FPV) P2 promoter sequence (see
CC   Q41013 and Q41018-20), multiple cloning sites (see Q41015-16), the FPV
CC   thymidine kinase (tk) gene and the tk gene intergenic region (see also
CC   Q41011-12). The P2 promoter belongs to one of the strongest natural
CC   promoters in vaccina virus (VV) infected cells. These plasmids may be
CC   used to express foreign proteins which require post-translational
CC   modification eg. Factors II, V, VII, VIII, IX, X, XI, XII, XIII,
CC   proteins C and S, von Willebrand factor, plasminogen and derivatives,
CC   apolipo proteins, eg. apoAI and apoAII, and viral antigens, eg.
CC   hepatitis B- antigens and the antigens of hepatitis C and E viruses,
CC   of tickborne encephalitis (TBE) virus, antigens of HIV, HSV and whole
CC   or partial sequences of antigens which cause pertussis, infectious
CC   bronchitis, tetanus, malaria, Marck's disease and Newcastle disease,
CC   these antigens being useful as vaccines.
SQ   Sequence   46 BP;    9 A;    4 C;    9 G;    24 T;


  Query Match            36.0%;  Score 27;  DB 1;  Length 46;
  Best Local Similarity  85.7%;  Pred. No. 6.5;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
          |||| |||||||||||||| |||||||  ||| ||
Db     11 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 45


RESULT  11
Q41025
ID   Q41025 standard; DNA; 62 BP.
AC   Q41025;
DT   29-JUL-1993  (first entry)
DE   P11 promoter mutant sart.
KW   Fowl pox virus; FPV; P11; promoter; vaccina virus; mutation; apoAI;
KW   multiple cloning site; plasmid; thymidine kinase; tk; Marck's disease;
KW   intergenic region; post-translational modification; factor II; apoAII;
KW   factor V; factor VII; factor VIII; factor IX; tick-borne encephalitis;
KW   factor X; factor XI; factor XII; factor XIII; protein C; protein S; VV;
KW   von Willebrand factor; hepatitis B; hepatitis C; hepatitis E; malaria;
KW   plasminogen; apolipo protein; viral antigen; infectious bronchitis;
KW   TBE; tetanus; HIV; HSV; pertussis; Newcastle disease; vaccine; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   promoter        34. .38
```

```
FT                  /*tag= a
FT                  /note= "Vaccinia virus late promoter consensus
FT                  sequence"
PN   AU9219580-A.
PD   04-MAR-1993.
PF   10-JUL-1992; 019580.
PR   26-AUG-1991; EP-114300.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Scheiflinger F;
DR   WPI; 93-126452/16.
PT   Prepn. of recombinant fowl pox virus for e.g. expression of
PT   proteins - by inserting foreign DNA into enlarged inter-genic
PT   region so FPV TK-gene remains intact and codes for entire
PT   thymidine kinase
PS   Disclosure; Fig 18; 92pp; English.
CC   The sequences given in Q41021-25 represent mutations of the vaccinia
CC   virus (VV) P11 promoter sequence which were included in the expression
CC   plasmids of the invention.  The sequences were unclear when given in
CC   the specification and so represent the "best guess" of sequence. These
CC   plasmids further comprised multiple cloning sites (see Q41015-16), the
CC   FPV thymidine kinase (tk) gene and the tk gene intergenic region (see
CC   also Q41011-12).  These plasmids may be used to express foreign
CC   proteins which require post-translational modification eg. Factors II,
CC   V, VII, VIII, IX, X, XI, XII, XIII, proteins C and S, von Willebrand
CC   factor, plasminogen and derivatives, apolipo proteins, eg. apoAI and
CC   apoAII, and viral antigens, eg. hepatitis B- antigens and the antigens
CC   of hepatitis C and E viruses, of tick-borne encephalitis (TBE) virus,
CC   antigens of HIV, HSV and whole or partial sequences of antigens which
CC   cause pertussis, infectious bronchitis, tetanus, malaria, Marck's
CC   disease and Newcastle disease, these antigens being useful as vaccines.
SQ   Sequence   62 BP;    14 A;    4 C;    11 G;    31 T;

  Query Match             36.0%;  Score 27;  DB 1;  Length 62;
  Best Local Similarity   85.7%;  Pred. No. 6.9;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      11 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 45


RESULT  12
Q40248
ID   Q40248 standard; DNA; 88 BP.
AC   Q40248;
DT   02-AUG-1993  (first entry)
DE   Sequence of plasmid pA1-S2 which comprises a first master cloning
DE   site comprised of unique sites of a second master cloning site of
DE   the vaccinia virus vector designated vdTK with a poxvirus promoter
DE   operatively linked to the first master cloning site.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Claim 63; Page 140; 206pp; English.
CC   Plasmids pA1-S2 and pA2-S2 comprise gene expression cassettes
CC   suitable for association of open reading frames already having a
CC   translation start codon with a synthetic poxvirus promoter (S2),
CC   prior to direct molecular transfer into vaccinia virus vector vdTK
CC   by forced cloning.  The S2 promoter is present in different
CC   orientations in the two plasmids.
SQ   Sequence   88 BP;    19 A;    13 C;    20 G;    36 T;

  Query Match             36.0%;  Score 27;  DB 1;  Length 88;
  Best Local Similarity   85.7%;  Pred. No. 7.4;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      31 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 65


RESULT  13
Q40249/c
ID   Q40249 standard; DNA; 92 BP.
AC   Q40249;
DT   02-AUG-1993  (first entry)
DE   Sequence of plasmid pA2-S2 which comprises a first master cloning
DE   site comprised of unique sites of a second master cloning site of
DE   the vaccinia virus vector designated vdTK with a poxvirus promoter
DE   operatively linked to the first master cloning site.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Claim 63; Page 140; 206pp; English.
CC   Plasmids pA1-S2 and pA2-S2 comprise gene expression cassettes
CC   suitable for association of open reading frames already having a
CC   translation start codon with a synthetic poxvirus promoter (S2),
CC   prior to direct molecular transfer into vaccinia virus vector vdTK
CC   by forced cloning.  The S2 promoter is present in different
CC   orientations in the two plasmids.
SQ   Sequence   92 BP;    37 A;    20 C;    14 G;    21 T;

  Query Match             36.0%;  Score 27;  DB 1;  Length 92;
  Best Local Similarity   85.7%;  Pred. No. 7.5;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      62 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 28


RESULT  14
Q40278/c
ID   Q40278 standard; DNA; 97 BP.
AC   Q40278;
DT   02-AUG-1993  (first entry)
DE   Sequence of synthetic linker consisting of the oligonucleotide
DE   P-artP(9) abd P-artP(10) used ot construct pN2gpt-S4.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Example; Page 153; 206pp; English.
CC   Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
```

```
CC   with a selective marker.  These plasmids were constructed by first
CC   making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC   gene driven by the vaccinia virus P7.5 promoter, flanked by
CC   several unique restriction sites including NotI.  For the
CC   construction of plasmid pN2gpt-S3A, the parental plasmid pN2-gptb
CC   was digested with PstI and ClaI and ligated with a synthetic
CC   linker consisting of the oligonucleotides P-artP(7) and P-artP(8).
CC   The synthetic promoter sequence of pN2gpt-S3A corresp. to the
CC   oligonucleotide P-artP(7).  For the construction of plasmid
CC   pN2gpt-S4, the plasmid pN2-gptb was digested with PstI and ClaI
CC   and ligated with an adaptor sequence consisting of the
CC   oligonucleotides P-artP(9) and P-artP(10).  The synthetic promoter
CC   sequence of pN2gpt-S4 corresp. to the oligonucleotide P-artP(9).
SQ   Sequence   97 BP;    32 A;    23 C;    22 G;    20 T;


  Query Match            35.5%;  Score 26.6;  DB 1;  Length 97;
  Best Local Similarity  66.7%;  Pred. No. 9.9;
  Matches   38;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGC 72
          |||| |||||||||||||| ||||||| ||    | |    |  ||||    | |||
Db     85 TTTTTTTTTTTTTTTTGGCATATAAATCGTTAACGAATTCCATGGCCCGGGAAGGC 29


RESULT  15
Q40250
ID   Q40250 standard; DNA; 134 BP.
AC   Q40250;
DT   02-AUG-1993  (first entry)
DE   Sequence of a DNA segment comprising a sequence-specific
DE   endonuclease cleavage site with a cleavage site for the bacterial
DE   restriction endonuclease NotI at each end and a selective marker
DE   gene and a second poxvirus promoter in plasmid pN2-gpt-S4.
KW   Plasmid; cloning; restriction site; ss.
OS   Synthetic.
PN   AU9221269-A.
PD   04-MAR-1993.
PF   25-AUG-1992; 021269.
PR   26-AUG-1991; US-750080.
PR   20-JUL-1992; US-914738.
PA   (IMMO ) IMMUNO AG.
PI   Dorner F,  Falkner FG,  Pfleiderer M,  Scheiflinger F;
DR   WPI; 93-126461/16.
PT   Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT   direct molecular cloning of modified DNA molecule contg.
PT   cytoplasmic DNA virus genome
PS   Claim 58; Page 141; 206pp; English.
CC   Plamid pN2 contains a DNA segment having a sequence-specific
CC   endonuclease cleavage site that is unique in the plasmid with a NotI
CC   site at each end. In this plasmid the DNA segment may further
CC   comprise a selective marker gene under transcriptional control of a
CC   poxvirus promoter, for instance pN2-gpta and pN2-gptb.  The DNA
CC   segment may further comprise a second poxvirus promoter operatively
CC   linked to a DNA sequence comprising a restriction endonuclease
CC   cleavage site, e.g. pN2gpt-S4.  pN2gpt-S4 are comprised of gene
CC   expresson cassettes suitable for association of an open reading
CC   frame having a translation start codon, with a synthetic
CC   promoter S4 prior to direct molecular transfer into a unique site in
CC   vaccinia virus vecotr vdTK.
SQ   Sequence   134 BP;    26 A;    34 C;    35 G;    39 T;


  Query Match            35.5%;  Score 26.6;  DB 1;  Length 134;
  Best Local Similarity  66.7%;  Pred. No. 11;
  Matches   38;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy     16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGC 72
          |||| |||||||||||||| ||||||| ||    | |    |  ||||    | |||
Db     29 TTTTTTTTTTTTTTTTGGCATATAAATCGTTAACGAATTCCATGGCCCGGGAAGGC 85
```

Search completed: May 29, 2000, 21:58:11
Job time: 38159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:08:49 ; Search time 621.83 Seconds
(without alignments)
15.678 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GGCCAAAAATTGAAATTTTA.....GCCCGGCCGCCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	27	36.0	46	1	US-08-232-463-40	Sequence 40, Appl
	2	27	36.0	63	1	US-08-232-463-47	Sequence 47, Appl
	3	27	36.0	88	1	US-07-750-080A-11	Sequence 11, Appl
c	4	27	36.0	92	1	US-07-750-080A-12	Sequence 12, Appl
c	5	27	36.0	9890	1	US-08-232-463-18	Sequence 18, Appl
c	6	26.6	35.5	97	1	US-07-750-080A-41	Sequence 41, Appl
	7	26.6	35.5	134	1	US-07-750-080A-14	Sequence 14, Appl
	8	26	34.7	1000	3	US-08-747-121-20	Sequence 20, Appl
	9	26	34.7	2002	3	US-08-747-121-1	Sequence 1, Appli
c	10	25.4	33.9	3471	3	US-08-715-568A-2	Sequence 2, Appli
	11	25.2	33.6	55	1	US-07-750-080A-38	Sequence 38, Appl
c	12	25.2	33.6	55	1	US-07-750-080A-39	Sequence 39, Appl
c	13	25.2	33.6	3095	5	US-08-434-000A-7	Sequence 7, Appli
c	14	25	33.3	3256	4	US-08-968-751-3	Sequence 3, Appli
c	15	24.8	33.1	2436	2	US-08-306-691B-16	Sequence 16, Appl
c	16	24.6	32.8	834	3	US-08-967-101-113	Sequence 113, App
c	17	24.6	32.8	834	4	US-08-592-541-113	Sequence 113, App
	18	24.6	32.8	2058	4	US-08-749-391-1	Sequence 1, Appli
c	19	24.6	32.8	3288	1	US-08-208-008C-3	Sequence 3, Appli
	20	23.8	31.7	42	4	US-08-484-575A-11	Sequence 11, Appl
	21	23.8	31.7	42	5	US-08-477-459-11	Sequence 11, Appl
	22	23.8	31.7	42	6	PCT-US94-01826A-11	Sequence 11, Appl
	23	23.8	31.7	42	6	PCT-US94-02252A-11	Sequence 11, Appl
	24	23.8	31.7	49	6	PCT-US96-07709-39	Sequence 39, Appl
c	25	23.8	31.7	49	6	PCT-US96-07709-40	Sequence 40, Appl
	26	23.8	31.7	49	6	PCT-US96-09848-19	Sequence 19, Appl
c	27	23.8	31.7	49	6	PCT-US96-09848-20	Sequence 20, Appl
	28	23.8	31.7	55	1	US-07-803-633A-3	Sequence 3, Appli
c	29	23.8	31.7	55	1	US-07-803-633A-4	Sequence 4, Appli
	30	23.8	31.7	55	3	US-08-525-742-13	Sequence 13, Appl
c	31	23.8	31.7	55	3	US-08-525-742-14	Sequence 14, Appl
	32	23.8	31.7	70	3	US-08-097-554A-43	Sequence 43, Appl
	33	23.8	31.7	70	5	US-08-480-640A-43	Sequence 43, Appl
	34	23.8	31.7	74	3	US-08-097-554A-44	Sequence 44, Appl
	35	23.8	31.7	74	5	US-08-480-640A-44	Sequence 44, Appl
	36	23.8	31.7	91	3	US-08-525-742-51	Sequence 51, Appl
c	37	23.8	31.7	93	1	US-07-750-080A-40	Sequence 40, Appl
c	38	23.8	31.7	93	3	US-08-525-742-28	Sequence 28, Appl
	39	23.8	31.7	95	3	US-08-525-742-29	Sequence 29, Appl
c	40	23.8	31.7	96	3	US-08-525-742-30	Sequence 30, Appl
	41	23.8	31.7	97	3	US-08-525-742-27	Sequence 27, Appl
	42	23.8	31.7	102	1	US-07-820-154A-10	Sequence 10, Appl
	43	23.8	31.7	102	3	US-08-097-554A-10	Sequence 10, Appl
	44	23.8	31.7	102	5	US-08-480-640A-10	Sequence 10, Appl
	45	23.8	31.7	102	6	PCT-US93-00324-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-40
; Sequence 40, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: sart(4)
US-08-232-463-40

```
  Query Match             36.0%;  Score 27;  DB 1;  Length 46;
  Best Local Similarity   85.7%;  Pred. No. 1.6;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      11 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 45


RESULT   2
US-08-232-463-47
; Sequence 47, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  sart
US-08-232-463-47


  Query Match             36.0%;  Score 27;  DB 1;  Length 63;
  Best Local Similarity   85.7%;  Pred. No. 1.8;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      11 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 45


RESULT   3
US-07-750-080A-11
; Sequence 11, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/750,080A
;      FILING DATE:  19910826
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  88 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  pA1-S2 (Fig. 4.5)
US-07-750-080A-11


  Query Match             36.0%;  Score 27;  DB 1;  Length 88;
  Best Local Similarity   85.7%;  Pred. No. 1.9;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      31 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 65


RESULT   4
US-07-750-080A-12/c
; Sequence 12, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/750,080A
;       FILING DATE:  19910826
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  92 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     IMMEDIATE SOURCE:
;       CLONE:  pA2-S2 (Fig. 4.5)
US-07-750-080A-12


  Query Match             36.0%;  Score 27;  DB 1;  Length 92;
  Best Local Similarity   85.7%;  Pred. No. 2;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||||||||||||| |||||||  ||| ||
Db      62 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 28


RESULT   5
US-08-232-463-18/c
; Sequence 18, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9890 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-sart
US-08-232-463-18


  Query Match             36.0%;  Score 27;  DB 1;  Length 9890;
  Best Local Similarity   85.7%;  Pred. No. 7.9;
  Matches   30;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
           |||| ||||,|||||||||| |||||||  ||| ||
Db    4326 TTTTTTTTTTTTTTTTTGGCATATAAATAGGCTGC 4292


RESULT   6
US-07-750-080A-41/c
; Sequence 41, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/750,080A
;      FILING DATE:  19910826
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  P-artP(10)
US-07-750-080A-41


  Query Match             35.5%;  Score 26.6;  DB 1;  Length 97;
  Best Local Similarity   66.7%;  Pred. No. 2.7;
  Matches   38;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGC 72
```

```
          |||| |||||||||||||| ||||||| ||    | |    |  ||||    | |||
Db      85 TTTTTTTTTTTTTTTTTGGCATATAAATCGTTAACGAATTCCATGGCCCGGGAAGGC 29


RESULT   7
US-07-750-080A-14
; Sequence 14, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/750,080A
;      FILING DATE:  19910826
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  134 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  pN2gpt-S4
US-07-750-080A-14


  Query Match             35.5%;  Score 26.6;  DB 1;  Length 134;
  Best Local Similarity   66.7%;  Pred. No. 2.9;
  Matches   38;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGC 72
           |||| |||||||||||||| ||||||| ||    | |    |  ||||    | |||
Db      29 TTTTTTTTTTTTTTTTTGGCATATAAATCGTTAACGAATTCCATGGCCCGGGAAGGC 85


RESULT   8
US-08-747-121-20
; Sequence 20, Application US/08747121
; Patent No. 5874290
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Gerald
;    APPLICANT:  Boynton, Alton
;    APPLICANT:  Sehgal, Anil
;    TITLE OF INVENTION:  NUCLEOTIDE AND AMINO ACID
;    TITLE OF INVENTION:  SEQUENCES OF A D2-2 GENE ASSOCIATED WITH
;    TITLE OF INVENTION:  BRAIN TUMORS AND METHODS BASED THEREON
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/747,121
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baldwin, Geraldine F
;      REGISTRATION NUMBER:  31,232
;      REFERENCE/DOCKET NUMBER:  8511-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)7909090
;      TELEFAX:  (212)8698864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1000 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-747-121-20


  Query Match             34.7%;  Score 26;  DB 3;  Length 1000;
  Best Local Similarity   70.0%;  Pred. No. 8.2;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGC 65
           |||| |||||||||||  || | | || | |   | | | || ||||||
Db      19 TTTTGTTTTTTTTTTTTTTAAGACAGAGAGTCGCTCGTCGCCCAGGCCGC 68


RESULT   9
US-08-747-121-1
; Sequence 1, Application US/08747121
; Patent No. 5874290
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Gerald
;    APPLICANT:  Boynton, Alton
;    APPLICANT:  Sehgal, Anil
;    TITLE OF INVENTION:  NUCLEOTIDE AND AMINO ACID
;    TITLE OF INVENTION:  SEQUENCES OF A D2-2 GENE ASSOCIATED WITH
;    TITLE OF INVENTION:  BRAIN TUMORS AND METHODS BASED THEREON
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/747,121
;      FILING DATE:  08-NOV-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baldwin, Geraldine F
;      REGISTRATION NUMBER:  31,232
;      REFERENCE/DOCKET NUMBER:  8511-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)7909090
;      TELEFAX:  (212)8698864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-747-121-1


  Query Match            34.7%;  Score 26;  DB 3;  Length 2002;
  Best Local Similarity  70.0%;  Pred. No. 10;
  Matches   35;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGC 65
           |||| |||||||||||||  || | | || | |   | | | || ||||||
Db    1021 TTTTGTTTTTTTTTTTTTTAAGACAGAGAGTCGCTCGTCGCCCAGGCCGC 1070


RESULT  10
US-08-715-568A-2/c
; Sequence 2, Application US/08715568A
; Patent No. 5856463
;  GENERAL INFORMATION:
;    APPLICANT:  Prydz, Hans Peter Blankenborg
;    APPLICANT:  Brede, Gaute
;    TITLE OF INVENTION:  PSKH-1 Ribozymes and Uses in Disease
;    TITLE OF INVENTION:  Treatment
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lerner, David, Littenberg, Krumholz & Mentlik
;      STREET:  600 South Avenue West
;      CITY:  Westfield
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07090-1497
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/715,568A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foley, Shawn P.
;      REFERENCE/DOCKET NUMBER:  FORSK 3.0-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-654-5000
;      TELEFAX:  908-654-7866
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3471 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-715-568A-2


  Query Match            33.9%;  Score 25.4;  DB 3;  Length 3471;
  Best Local Similarity  68.6%;  Pred. No. 18;
  Matches   35;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCC 66
           |||| |||||||||| |||||||  |  || |||  |   ||   ||| |||
Db    3471 TTTTTTTTTTTTTTTGTTGGAATGAATGGCAGCCCTCCCTGGAGAGGCAGCC 3421


RESULT  11
US-07-750-080A-38
; Sequence 38, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/750,080A
;      FILING DATE:  19910826
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  P-artP(11)
US-07-750-080A-38


  Query Match            33.6%;  Score 25.2;  DB 1;  Length 55;
  Best Local Similarity  90.0%;  Pred. No. 6.1;
  Matches   27;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCG 45
           |||| |||||||||||||| ||||||| ||
Db      25 TTTTTTTTTTTTTTTTTGGCATATAAATCG 54


RESULT  12
US-07-750-080A-39/c
; Sequence 39, Application US/07750080A
; Patent No. 5445953
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/750,080A
;      FILING DATE:  19910826
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  P-artP(12)
US-07-750-080A-39


  Query Match             33.6%;  Score 25.2;  DB 1;  Length 55;
  Best Local Similarity   90.0%;  Pred. No. 6.1;
  Matches   27;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCG 45
           |||| |||||||||||||| ||||||| ||
Db      35 TTTTTTTTTTTTTTTTTGGCATATAAATCG 6


RESULT  13
US-08-434-000A-7/c
; Sequence 7, Application US/08434000A
; Patent No. 6046037
;  GENERAL INFORMATION:
;    APPLICANT:  ANDREW C. HIATT, JULIAN
;    APPLICANT:  K.-C. MA, THOMAS LEHNER
;    TITLE OF INVENTION:  IMMUNOGLOBULINS CONTAINING PROTECTION
;    TITLE OF INVENTION:  PROTEINS IN PLANTS AND THEIR USES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/434,000A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:          1
;      APPLICATION NUMBER:  08/367,395
;      FILING DATE:  12/30/94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Guise, Jeffrey W.
;      REGISTRATION NUMBER:  34,613
;      REFERENCE/DOCKET NUMBER:  212/127
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 552-8400
;      TELEFAX:  (619) 552-0159
;      TELEX:  67-3510
;      TELEX:  SEQUENCE LISTING
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3095 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;      TOPOLOGY:  DESCRIPTION:  Mouse Polyimmunoglobulin Receptor
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  85....2400
US-08-434-000A-7


  Query Match             33.6%;  Score 25.2;  DB 5;  Length 3095;
  Best Local Similarity   71.7%;  Pred. No. 20;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGG 61
           |||| |||||||||| | ||| | | | ||   |||| | ||| ||
Db    3092 TTTTTTTTTTTTTTTCTTGAAGACAGATCGTAGGCCAAGAGCCAGG 3047


RESULT  14
US-08-968-751-3/c
; Sequence 3, Application US/08968751
; Patent No. 5948643
;  GENERAL INFORMATION:
;    APPLICANT:  Rubinfeld, Bonnee
;    APPLICANT:  Polakis, Paul G.
;    APPLICANT:  Ligenfelter, Carol
;    APPLICANT:  Vuong, Terilyn T.
;    TITLE OF INVENTION:  MODULATORS OF BRCA1 ACTIVITY
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ONYX Pharmaceuticals, Inc.
;      STREET:  3031 Research Drive
;      CITY:  Richmond
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94806
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/968,751
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Giotta, Gregory
;      REGISTRATION NUMBER:  32,028
;      REFERENCE/DOCKET NUMBER:  ONYX1024 GG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 262-8710
;      TELEFAX:  (510) 222-9758
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3256 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  34..2541
US-08-968-751-3


  Query Match             33.3%;  Score 25;  DB 4;  Length 3256;
  Best Local Similarity   84.8%;  Pred. No. 24;
  Matches   28;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      10 TTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
           | ||  |||| |||||||||||||| |||||||
Db    3237 TCGAGTTTTTTTTTTTTTTTTTTGGTATATAAA 3205


RESULT  15
US-08-306-691B-16/c
; Sequence 16, Application US/08306691B
; Patent No. 5734039
;  GENERAL INFORMATION:
;     APPLICANT:  Calabretta, Bruno
;     APPLICANT:  Skorski, Tomasz
;     TITLE OF INVENTION:  ANTISENSE
;     TITLE OF INVENTION:  OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
;     NUMBER OF SEQUENCES:  55
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seidel, Gonda, Lavorgna & Monaco, P.C.
;       STREET:  Two Penn Center, Suite 1800
;       CITY:  Philadelphia
;       STATE:  Pennsylvania
;       COUNTRY:  U.S.A.
;       ZIP:  19102
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/306,691B
;       FILING DATE:  September 15, 1994
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monaco, Daniel A.
;       REGISTRATION NUMBER:  30,480
;       REFERENCE/DOCKET NUMBER:  8321-8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (215) 568-8383
;       TELEFAX:  (215) 568-5549
;       TELEX:  No. 5734039e
;  INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2436 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-306-691B-16


  Query Match             33.1%;  Score 24.8;  DB 2;  Length 2436;
  Best Local Similarity   67.3%;  Pred. No. 25;
  Matches   35;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCA 52
           | |  |||||| |   || |||||| | |||  |||| || ||| |  ||||
Db    1894 GACTCAAAATTCAGTATTCATTTTTGTATTTATGAATCTAGAGCAGATGCCA 1843
```

```
Search completed:  May 29, 2000, 22:08:54
Job time:  38735 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:19 ; Search time 2276.24 Seconds
(without alignments)
147.011 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GGCCAAAAATTGAAATTTTA.....GCCCGGCCGCCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	75	100.0	75	28	US-08-935-377-3	Sequence 3, Appli
2	59	78.7	59	28	US-08-935-377-25	Sequence 25, Appl

c	3	59	78.7	59	28	US-08-935-377-26	Sequence 26, Appl
	4	42.4	56.5	77	28	US-08-935-377-32	Sequence 32, Appl
	5	40	53.3	84	12	US-08-322-358-8	Sequence 8, Appli
c	6	40	53.3	86	12	US-08-322-358-9	Sequence 9, Appli
	7	40	53.3	4723	12	US-08-322-358-10	Sequence 10, Appl
c	8	39.6	52.8	6474	3	US-07-914-738B-66	Sequence 66, Appl
c	9	39.6	52.8	6474	13	US-08-358-928-66	Sequence 66, Appl
c	10	39.6	52.8	6474	18	US-08-651-472-66	Sequence 66, Appl
c	11	38.8	51.7	5532	3	US-07-914-738B-72	Sequence 72, Appl
c	12	38.8	51.7	5532	13	US-08-358-928-72	Sequence 72, Appl
c	13	38.8	51.7	5532	18	US-08-651-472-72	Sequence 72, Appl
	14	38.4	51.2	49	3	US-07-914-738B-70	Sequence 70, Appl
	15	38.4	51.2	49	13	US-08-358-928-70	Sequence 70, Appl
	16	38.4	51.2	49	18	US-08-651-472-70	Sequence 70, Appl
c	17	38.4	51.2	77	28	US-08-935-377-33	Sequence 33, Appl
c	18	38.4	51.2	6811	3	US-07-914-738B-67	Sequence 67, Appl
c	19	38.4	51.2	6811	13	US-08-358-928-67	Sequence 67, Appl
c	20	38.4	51.2	6811	18	US-08-651-472-67	Sequence 67, Appl
	21	38.4	51.2	7450	12	US-08-322-358-11	Sequence 11, Appl
	22	38.4	51.2	9803	12	US-08-322-358-12	Sequence 12, Appl
	23	38	50.7	51	3	US-07-914-738B-56	Sequence 56, Appl
	24	38	50.7	51	13	US-08-358-928-56	Sequence 56, Appl
	25	38	50.7	51	18	US-08-651-472-56	Sequence 56, Appl
c	26	38	50.7	53	3	US-07-914-738B-57	Sequence 57, Appl
c	27	38	50.7	53	13	US-08-358-928-57	Sequence 57, Appl
c	28	38	50.7	53	18	US-08-651-472-57	Sequence 57, Appl
c	29	38	50.7	3878	3	US-07-914-738B-65	Sequence 65, Appl
c	30	38	50.7	3878	13	US-08-358-928-65	Sequence 65, Appl
c	31	38	50.7	3878	18	US-08-651-472-65	Sequence 65, Appl
	32	37	49.3	69	28	US-08-935-377-1	Sequence 1, Appli
	33	31	41.3	258	46	US-09-288-687-4416	Sequence 4416, Ap
	34	30.6	40.8	255	46	US-09-288-687-4576	Sequence 4576, Ap
c	35	30.6	40.8	525	95	US-60-182-569-372	Sequence 372, App
c	36	30.6	40.8	640	104	US-60-195-135-32	Sequence 32, Appl
c	37	30.4	40.5	501	8	US-08-196-362A-1517	Sequence 1517, Ap
c	38	30.4	40.5	501	8	US-08-196-362D-1517	Sequence 1517, Ap
c	39	30.4	40.5	570	54	US-09-401-645-4227	Sequence 4227, Ap
	40	30	40.0	403	47	US-09-293-972-33060	Sequence 33060, A
	41	30	40.0	403	57	US-09-515-126-2413	Sequence 2413, Ap
c	42	29.8	39.7	442	47	US-09-293-972-9281	Sequence 9281, Ap
c	43	29.6	39.5	419	51	US-09-353-690-1777	Sequence 1777, Ap
	44	29.6	39.5	1394	86	US-60-144-351-4913	Sequence 4913, Ap
c	45	29.4	39.2	349	49	US-09-321-214-18207	Sequence 18207, A

ALIGNMENTS

```
RESULT   1
US-08-935-377-3
; Sequence 3, Application US/08935377
;  GENERAL INFORMATION:
;     APPLICANT:  Zauderer, Maurice
;     TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;     TITLE OF INVENTION:  Vaccines Based Thereon
;     NUMBER OF SEQUENCES:  37
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;       STREET:  1100 New York Avenue, N.W., Suite 600
;       CITY:  Washington
;       STATE:  D. C.
;       COUNTRY:  USA
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/935,377
;       FILING DATE:  22-SEP-1997
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Steffe, Eric K
;       REGISTRATION NUMBER:  36,688
;       REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  75 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  52..75
US-08-935-377-3


  Query Match           100.0%;  Score 75;  DB 28;  Length 75;
  Best Local Similarity  100.0%;  Pred. No. 5.3e-09;
  Matches   75;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60

Qy      61 GCCGCCAACGGCGGA 75
           |||||||||||||||
Db      61 GCCGCCAACGGCGGA 75


RESULT   2
US-08-935-377-25
; Sequence 25, Application US/08935377
;  GENERAL INFORMATION:
;     APPLICANT:  Zauderer, Maurice
;     TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;     TITLE OF INVENTION:  Vaccines Based Thereon
;     NUMBER OF SEQUENCES:  37
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;       STREET:  1100 New York Avenue, N.W., Suite 600
;       CITY:  Washington
;       STATE:  D. C.
;       COUNTRY:  USA
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/935,377
;       FILING DATE:  22-SEP-1997
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Steffe, Eric K
;       REGISTRATION NUMBER:  36,688
;       REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  59 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-935-377-25
```

```
  Query Match             78.7%;  Score 59;  DB 28;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-05;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCC 59
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCC 59


RESULT   3
US-08-935-377-26/c
; Sequence 26, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-26


  Query Match             78.7%;  Score 59;  DB 28;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-05;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCC 63
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCC 1


RESULT   4
US-08-935-377-32
; Sequence 32, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-32


  Query Match             56.5%;  Score 42.4;  DB 28;  Length 77;
  Best Local Similarity   97.7%;  Pred. No. 0.31;
  Matches   43;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTGGAATATAAAGC 44
          |||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTGGAATATAAACC 44


RESULT   5
US-08-322-358-8
; Sequence 8, Application US/08322358
;  GENERAL INFORMATION:
;    APPLICANT:  FALKNER, Falko-Gunter
;    APPLICANT:  BODEMER, Walter
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  MOSS, Bernard
;    TITLE OF INVENTION:  PRODUCTION OF ISOLATED PROTEINACEOUS
;    TITLE OF INVENTION:  MATERIALS USING RECOMBINANT AVIPOX VIRUS VECTORS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/322,358
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/882,768
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/734,741
;      FILING DATE:  23-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30357/108 KIST
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  84 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     IMMEDIATE SOURCE:
;       CLONE:  Oligonucleotide oselP2
US-08-322-358-8


  Query Match             53.3%;  Score 40;  DB 12;  Length 84;
  Best Local Similarity   82.1%;  Pred. No. 1.2;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           | ||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db       9 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 64


RESULT   6
US-08-322-358-9/c
; Sequence 9, Application US/08322358
;  GENERAL INFORMATION:
;     APPLICANT:  FALKNER, Falko-Gunter
;     APPLICANT:  BODEMER, Walter
;     APPLICANT:  DORNER, Friedrich
;     APPLICANT:  MOSS, Bernard
;     TITLE OF INVENTION:  PRODUCTION OF ISOLATED PROTEINACEOUS
;     TITLE OF INVENTION:  MATERIALS USING RECOMBINANT AVIPOX VIRUS VECTORS
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley & Lardner
;       STREET:  1800 Diagonal Road, Suite 500
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/322,358
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/882,768
;       FILING DATE:
;       APPLICATION NUMBER:  US 07/734,741
;       FILING DATE:  23-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30357/108 KIST
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  86 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     IMMEDIATE SOURCE:
;       CLONE:  Oligonucleotide oselP2
US-08-322-358-9


  Query Match             53.3%;  Score 40;  DB 12;  Length 86;
  Best Local Similarity   82.1%;  Pred. No. 1.2;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           | ||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db      80 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 25


RESULT   7
US-08-322-358-10
; Sequence 10, Application US/08322358
;  GENERAL INFORMATION:
;     APPLICANT:  FALKNER, Falko-Gunter
;     APPLICANT:  BODEMER, Walter
;     APPLICANT:  DORNER, Friedrich
;     APPLICANT:  MOSS, Bernard
;     TITLE OF INVENTION:  PRODUCTION OF ISOLATED PROTEINACEOUS
;     TITLE OF INVENTION:  MATERIALS USING RECOMBINANT AVIPOX VIRUS VECTORS
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley & Lardner
;       STREET:  1800 Diagonal Road, Suite 500
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/322,358
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/882,768
;       FILING DATE:
;       APPLICATION NUMBER:  US 07/734,741
;       FILING DATE:  23-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30357/108 KIST
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4723 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     IMMEDIATE SOURCE:
;       CLONE:  Plasmid pTKgpt-selP (4723bp)
US-08-322-358-10


  Query Match             53.3%;  Score 40;  DB 12;  Length 4723;
  Best Local Similarity   82.1%;  Pred. No. 2.3;
  Matches   46;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           | ||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db    1524 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 1579
```

```
RESULT   8
US-07-914-738B-66/c
; Sequence 66, Application US/07914738B
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    APPLICANT:  PFLEIDERER, M.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/914,738B
;      FILING DATE:  19920720
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  66:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6474 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  pselP-gp160MN
US-07-914-738B-66


  Query Match             52.8%;  Score 39.6;  DB 3;  Length 6474;
  Best Local Similarity   83.3%;  Pred. No. 3;
  Matches   45;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           ||||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db    3964 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 3911


RESULT   9
US-08-358-928-66/c
; Sequence 66, Application US/08358928
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  SCHEIFLINGER, Friedrich
;    APPLICANT:  FALKNER, Falko Gunter
;    APPLICANT:  PFLEIDERER, Michael
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;    TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;    TITLE OF INVENTION:  (HIV-1) ANTIGENS
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/358,928
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/914,738
;      FILING DATE:  20-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904135
;  INFORMATION FOR SEQ ID NO:  66:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6474 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  pselP-gp160MN
US-08-358-928-66


  Query Match             52.8%;  Score 39.6;  DB 13;  Length 6474;
  Best Local Similarity   83.3%;  Pred. No. 3;
  Matches   45;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           ||||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db    3964 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 3911


RESULT  10
US-08-651-472-66/c
; Sequence 66, Application US/08651472
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  SCHEIFLINGER, Friedrich
;    APPLICANT:  FALKNER, Falko Gunter
;    APPLICANT:  PFLEIDERER, Michael
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;    TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;    TITLE OF INVENTION:  (HIV-1) ANTIGENS
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/651,472
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/914,738
;       FILING DATE:  20-JUL-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/750,080
;       FILING DATE:  26-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)672-5300
;       TELEFAX:  (202)672-5399
;       TELEX:  904136
;   INFORMATION FOR SEQ ID NO:  66:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6474 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid;
;       DESCRIPTION:  Synthetic DNA oligonucleotide
;     IMMEDIATE SOURCE:
;       CLONE:  pselP-gp160MN
US-08-651-472-66


  Query Match             52.8%;  Score 39.6;  DB 18;  Length 6474;
  Best Local Similarity   83.3%;  Pred. No. 3;
  Matches   45;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCC 58
           |||||||||||||||||||||||||||||||||||||   | |  |   ||||
Db    3964 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGCC 3911


RESULT  11
US-07-914-738B-72/c
; Sequence 72, Application US/07914738B
;  GENERAL INFORMATION:
;     APPLICANT:  DORNER, F.
;     APPLICANT:  SCHEIFLINGER, F.
;     APPLICANT:  FALKNER, F. G.
;     APPLICANT:  PFLEIDERER, M.
;     TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;     TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;     NUMBER OF SEQUENCES:  84
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley & Lardner
;       STREET:  1800 Diagonal Road, Suite 500
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/914,738B
;       FILING DATE:  19920720
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/750,080
;       FILING DATE:  26-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  72:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5532 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid;
;       DESCRIPTION:  Synthetic DNA oligonucleotide
;     IMMEDIATE SOURCE:
;       CLONE:  pN2gpta-FIX
US-07-914-738B-72


  Query Match             51.7%;  Score 38.8;  DB 3;  Length 5532;
  Best Local Similarity   79.3%;  Pred. No. 4.6;
  Matches   46;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
           | |||||||||||||||||||||||||||||||||||||   | |  |   ||  ||
Db    3709 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGAGCG 3652


RESULT  12
US-08-358-928-72/c
; Sequence 72, Application US/08358928
;  GENERAL INFORMATION:
;     APPLICANT:  DORNER, Friedrich
;     APPLICANT:  SCHEIFLINGER, Friedrich
;     APPLICANT:  FALKNER, Falko Gunter
;     APPLICANT:  PFLEIDERER, Michael
;     TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;     TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;     TITLE OF INVENTION:  (HIV-1) ANTIGENS
;     NUMBER OF SEQUENCES:  95
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley & Lardner
;       STREET:  3000 K Street, N.W., Suite 500
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20007-5109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/358,928
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/914,738
;       FILING DATE:  20-JUL-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/750,080
;       FILING DATE:  26-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)672-5300
;       TELEFAX:  (202)672-5399
;       TELEX:  904136
;   INFORMATION FOR SEQ ID NO:  72:
;     SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  5532 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  pN2gpta-FIX
US-08-358-928-72


  Query Match             51.7%;  Score 38.8;  DB 13;  Length 5532;
  Best Local Similarity   79.3%;  Pred. No. 4.6;
  Matches   46;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
          | ||||||||||||||||||||||||||||||||||||||   | |  |   ||  ||
Db   3709 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGAGCG 3652


RESULT  13
US-08-651-472-72/c
; Sequence 72, Application US/08651472
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  SCHEIFLINGER, Friedrich
;    APPLICANT:  FALKNER, Falko Gunter
;    APPLICANT:  PFLEIDERER, Michael
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;    TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;    TITLE OF INVENTION:  (HIV-1) ANTIGENS
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/651,472
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/914,738
;      FILING DATE:  20-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5532 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  pN2gpta-FIX
US-08-651-472-72


  Query Match             51.7%;  Score 38.8;  DB 18;  Length 5532;
  Best Local Similarity   79.3%;  Pred. No. 4.6;
  Matches   46;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
          | ||||||||||||||||||||||||||||||||||||||   | |  |   ||  ||
Db   3709 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAATAAGGCCTCCATGGAGCG 3652


RESULT  14
US-07-914-738B-70
; Sequence 70, Application US/07914738B
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    APPLICANT:  PFLEIDERER, M.
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF A MODIFIED
;    TITLE OF INVENTION:  EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/914,738B
;      FILING DATE:  19920720
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/106 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  selP promoter
US-07-914-738B-70


  Query Match             51.2%;  Score 38.4;  DB 3;  Length 49;
  Best Local Similarity   97.5%;  Pred. No. 2.6;
  Matches   39;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
          | |||||||||||||||||||||||||||||||||||||||
Db      9 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 48


RESULT  15
US-08-358-928-70
```

```
; Sequence 70, Application US/08358928
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, Friedrich
;    APPLICANT:  SCHEIFLINGER, Friedrich
;    APPLICANT:  FALKNER, Falko Gunter
;    APPLICANT:  PFLEIDERER, Michael
;    TITLE OF INVENTION:  DIRECT MOLECULAR CLONING OF CHIMERIC
;    TITLE OF INVENTION:  VIRUSES CONTAINING HUMAN IMMUNODEFCIENCY VIRUS TYPE 1
;    TITLE OF INVENTION:  (HIV-1) ANTIGENS
;    NUMBER OF SEQUENCES:  95
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/358,928
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/914,738
;      FILING DATE:  20-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/750,080
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/166/IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid;
;      DESCRIPTION:  Synthetic DNA oligonucleotide
;    IMMEDIATE SOURCE:
;      CLONE:  selP promoter
US-08-358-928-70


  Query Match             51.2%;  Score 38.4;  DB 13;  Length 49;
  Best Local Similarity   97.5%;  Pred. No. 2.6;
  Matches   39;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
           | ||||||||||||||||||||||||||||||||||||||
Db       9 CTAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 48


Search completed: May 30, 2000, 09:48:27
Job time: 60177 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:10 ; Search time 2192.43 Seconds
(without alignments)
138.655 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GGCCAAAAATTGAAATTTTA.....GCCCGGCCGCCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	35.4	47.2	1201	83	CNS01651	AL106303 Drosophil
2	33.4	44.5	158	79	AW297017	AW297017 UI-H-BW0-
c 3	30.6	40.8	1101	82	CNS00YW5	AL096911 Drosophil
4	30.4	40.5	909	85	AQ743767	AQ743767 HS_5501_A
5	30	40.0	607	48	AI570555	AI570555 tm78a09.x
c 6	30	40.0	872	90	AQ752434	AQ752434 HS_5566_B
c 7	29.8	39.7	1404	63	AI964281	AI964281 EST269395
c 8	29.4	39.2	468	79	AW297722	AW297722 UI-H-BW0-
c 9	29.2	38.9	230	32	AA339067	AA339067 EST44123
c 10	29.2	38.9	644	62	AI902252	AI902252 QV-BT003-
11	29.2	38.9	1024	82	CNS00F3L	AL069994 Drosophil
12	29	38.7	368	64	AW026618	AW026618 wv45c10.x
c 13	28.8	38.4	378	28	AA123454	AA123454 mp94a05.r
c 14	28.8	38.4	423	53	AI683617	AI683617 tw52g05.x
c 15	28.8	38.4	425	26	W30577	W30577 mc28a08.r1
16	28.8	38.4	464	69	AW125685	AW125685 UI-M-BH2.
17	28.8	38.4	553	70	AW155524	AW155524 mgie0030P
18	28.8	38.4	805	90	AQ782146	AQ782146 HS_3174_B
c 19	28.8	38.4	1062	82	CNS00CAA	AL058926 Drosophil
c 20	28.6	38.1	194	54	AU073906	AU073906 AU073906
c 21	28.6	38.1	372	44	AI285735	AI285735 qu81d09.x
22	28.6	38.1	953	82	CNS002MH	AL097859 Drosophil
23	28.6	38.1	1201	83	CNS0162C	AL106206 Drosophil
c 24	28.4	37.9	261	73	AW168602	AW168602 xi90f01.x
25	28.4	37.9	421	44	AI253335	AI253335 ap42a11.x
26	28.4	37.9	489	44	AI267634	AI267634 aq92d04.x
27	28.4	37.9	710	90	AQ780818	AQ780818 HS_3104_B
28	28.4	37.9	1101	83	CNS016YN	AL107369 Drosophil
c 29	28.2	37.6	184	60	AI783861	AI783861 tr29c08.x
c 30	28.2	37.6	409	54	AI738854	AI738854 tr29h02.x
31	28.2	37.6	498	34	AA536039	AA536039 nj81c08.s
c 32	28.2	37.6	720	36	AA660391	AA660391 00266 MtR
33	28.2	37.6	1101	82	CNS0001Y	AL050448 Drosophil
c 34	28.2	37.6	1101	82	CNS001QN	AL061241 Drosophil
35	28	37.3	335	53	AI679321	AI679321 tu73b11.x
36	28	37.3	445	47	AI493105	AI493105 qz49g04.x
37	28	37.3	648	52	AI667092	AI667092 fc37e11.x
38	28	37.3	847	84	B20040	B20040 T7B22-T7.2
39	28	37.3	1101	82	CNS00HZN	AL074548 Drosophil
40	27.8	37.1	224	64	AW074018	AW074018 xb06d08.x
41	27.8	37.1	248	46	AI445994	AI445994 tj32c09.x
c 42	27.8	37.1	389	38	AA773220	AA773220 af78e04.r
43	27.8	37.1	522	52	AI629893	AI629893 486041F08
44	27.8	37.1	531	40	AA947860	AA947860 ok21g01.s
45	27.8	37.1	602	105	AQ320590	AQ320590 RPCI11-96

ALIGNMENTS

RESULT 1
CNS01651/c
LOCUS CNS01651 1201 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC BACN15M07 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL106303
VERSION AL106303.1 GI:5621247
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES Location/Qualifiers
source 1. .1201
/organism="Drosophila melanogaster"
/plasmid="pBeloBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN15M07"
/note="end : T7"
BASE COUNT 286 a 290 c 192 g 283 t 150 others
ORIGIN

Query Match 47.2%; Score 35.4; DB 83; Length 1201;
Best Local Similarity 47.5%; Pred. No. 7.8;
Matches 29; Conservative 22; Mismatches 10; Indels 0; Gaps 0;

```
Qy       9 ATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAA 68
           :::|   |M| ||||||||:::::: |:| |||| |  ::|::|||::| :|:::||::
Db     960 RWKGTTTTTTTTTTTTTTTTTKKKKDWTTWTTAAGCTGKTSSCWKGGGHYCCSCSSCCMM 901

Qy      69 C 69
           :
Db     900 M 900
```

RESULT 2
AW297017
LOCUS AW297017 158 bp mRNA EST 16-JAN-2000
DEFINITION UI-H-BW0-ajf-d-09-0-UI.s1 NCI_CGAP_Sub6 Homo sapiens cDNA clone IMAGE:2731457 3', mRNA sequence.
ACCESSION AW297017
VERSION AW297017.1 GI:6703653
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 158)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133426.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html The following repetitive elements were found in this cDNA sequence: 99-142,
>GC_rich#Low_complexity
Seq primer: M13 Forward
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .158
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2731457"
/clone_lib="NCI_CGAP_Sub6"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub6 is a subtracted library derived from BW, which consists of a mixture of four normalized libraries: NCI_CGAP_Brn50, NCI_CGAP_Lu13, NCI_CGAP_Ov18, GBC1 . The NCI_CGAP_Sub6 library had 7 million recombinants. A single-stranded DNA preparation of BW was used as a tracer in a subtractive hybridization with a driver comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775,1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342,3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,1144584-1145351). (50% of the driver population), plus a pool of 3,840 arrayed clones from NCI_CGAP_Sub1 (IMAGE CloneIDs 2708616-2710535) and NCI_CGAP_Sub2 (IMAGE CloneIDs 2710536-2712455) (20% of the driver population), plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE CloneIDs 2712456-2723591) (30% of the driver population). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI_CGAP-Lu13
TAG_TISSUE=lung
TAG_SEQ=GCCGG"
BASE COUNT 17 a 62 c 43 g 36 t
ORIGIN

Query Match 44.5%; Score 33.4; DB 79; Length 158;
Best Local Similarity 72.9%; Pred. No. 49;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 16 TTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACGGCGG 74
Db 1 TTTTTTTTTTTTTTTTTGGTAGGAACAGCGTCTTTAATGGGCCCCGCCCCCCACGCCCG 59

RESULT 3
CNS00YW5/c
LOCUS CNS00YW5 1101 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC BACN01A18 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL096911
VERSION AL096911.1 GI:5608522
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/plasmid="pBeloBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN01A18"
/note="end : T7"
BASE COUNT 270 a 179 c 175 g 330 t 147 others
ORIGIN

Query Match 40.8%; Score 30.6; DB 82; Length 1101;
Best Local Similarity 49.2%; Pred. No. 1.2e+02;
Matches 32; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Qy 2 GCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGG 61
Db 1000 SCCAAAWAATTTWAWAWWTTTTTTTTTTWTRGRGADGVVVSVACCCCCCCCTGRCCCGG 941

Qy 62 CCGCC 66
Db 940 TTGMC 936

RESULT 4
AQ743767
LOCUS AQ743767 909 bp DNA GSS 16-JUL-1999
DEFINITION HS_5501_A1_C11_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1077 Col=21 Row=E, genomic survey sequence.
ACCESSION AQ743767
VERSION AQ743767.1 GI:5521289
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 1077 row: E column: 21
Seq primer: T7
Class: BAC ends
High quality sequence stop: 909.
FEATURES Location/Qualifiers
source 1. .909
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=1077 Col=21 Row=E"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Genomic sequence of BAC ends"
BASE COUNT 217 a 153 c 315 g 190 t 34 others
ORIGIN

```
  Query Match             40.5%;  Score 30.4;  DB 85;  Length 909;
  Best Local Similarity   63.9%;  Pred. No. 1.5e+02;
  Matches   46;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGC 62
          |||||||   |||||||  ||||||||||||   | | |||    || ||     |||  |
Db    157 CCAAAAAGAAAAATTTTTTTTTTTTTTTTTTTTAAAAAAAAAACCCCCCCCCCCCCCCC 216

Qy     63 CGCCAACGGCGG 74
          | ||   || ||
Db    217 CCCCGGGGGGGG 228


RESULT   5
AI570555
LOCUS       AI570555      607 bp    mRNA            EST       14-APR-1999
DEFINITION  tm78a09.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2164216 3'
            similar to WP:Y57G11C.13 CE14950 ADP-RIBOSYLATION FACTOR ;, mRNA
            sequence.
ACCESSION   AI570555
VERSION     AI570555.1  GI:4533929
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 607)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     On Mar 10, 1998 this sequence version replaced gi:2949517.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1231   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 441.
FEATURES             Location/Qualifiers
     source          1. .607
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2164216"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      144 a    151 c    127 g    184 t      1 others
ORIGIN


  Query Match             40.0%;  Score 30;  DB 48;  Length 607;
  Best Local Similarity   67.7%;  Pred. No. 2.1e+02;
  Matches   42;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy     11 TGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCGCCAACG 70
          || |  ||| |||||||||||||   | || || |   ||||| | ||| || | || ||
Db     48 TGTAGGTTTTTTTTTTTTTTTTCATACATGAAACTCACGCCAGGTGCCGGGACCCCCACA 107

Qy     71 GC 72
          ||
Db    108 GC 109


RESULT   6
AQ752434/c
LOCUS       AQ752434      872 bp    DNA             GSS       19-JUL-1999
DEFINITION  HS_5566_B2_E08_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1142 Col=16 Row=J, genomic survey sequence.
ACCESSION   AQ752434
VERSION     AQ752434.1  GI:5539592
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 872)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     On Feb 19, 1999 this sequence version replaced gi:4145341.
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1142  row: J  column: 16
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 872.
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1142 Col=16 Row=J"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Genomic sequence of BAC ends"
BASE COUNT      255 a    212 c    154 g    237 t     14 others
ORIGIN


  Query Match             40.0%;  Score 30;  DB 90;  Length 872;
  Best Local Similarity   78.3%;  Pred. No. 1.9e+02;
  Matches   36;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      5 AAAAATTGAAATTTTATTTTTTTTTTTTTGGAATATAAAGCGGCCGC 50
          |||||   ||| ||||||||||||||||  ||||  ||| || |||
Db    871 AAAAAAAAAAAATTTATTTTTTTTTTTTTTAATAATAAGAGGGCGC 826


RESULT   7
AI964281/c
```

```
LOCUS       AI964281     1404 bp    mRNA            EST       20-AUG-1999
DEFINITION  EST269395 zebrafish, Steve Ekker Danio rerio cDNA clone RZBDA89,
            mRNA sequence.
ACCESSION   AI964281
VERSION     AI964281.1  GI:5756994
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
            Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
            Cyprinoidea; Cyprinidae; Rasborinae; Danio.
REFERENCE   1  (bases 1 to 1404)
  AUTHORS   Quackenbush,J., Lee,N., Tomb,J.-F. and Ekker,S.
  TITLE     Generation of ESTs from zebrafish
  JOURNAL   Unpublished (1999)
COMMENT     On May 18, 1998 this sequence version replaced gi:3136813.
            Other_ESTs: TC1242
            Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            3 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="RZBDA89"
                     /clone_lib="zebrafish, Steve Ekker"
                     /tissue_type="whole body"
                     /dev_stage="maternal"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pAD-GAL4; Site_1: EcoR1; Site_2: Xho1; This
                     library was constructed by directed cloning into the
                     EcoRI-XhoI sites of pAD-GAL4.Estimated Average insert size
                      approximately 1.2kb based on 15 clones, estimated titer
                     of  excised phagemid library = 3.78 x 106 cfu/ml"
BASE COUNT      397 a    307 c    267 g    421 t     12 others
ORIGIN


  Query Match             39.7%;  Score 29.8;  DB 63;  Length 1404;
  Best Local Similarity   66.2%;  Pred. No. 1.8e+02;
  Matches   43;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       3 CCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGC 62
           ||||||||| ||||||| |||||||||| |     |||  | ||   ||    |   |||
Db    1351 CCAAAAATTAAAATTTTTTTTTTTTTTTTATTACCAATACCGAGGGGTCCCCCAGGAGGGC 1292

Qy      63 CGCCA 67
           | |||
Db    1291 CCCCA 1287


RESULT   8
AW297722/c
LOCUS       AW297722      468 bp    mRNA            EST       16-JAN-2000
DEFINITION  UI-H-BW0-ajh-h-05-0-UI.s1 NCI_CGAP_Sub6 Homo sapiens cDNA clone
            IMAGE:2732001 3', mRNA sequence.
ACCESSION   AW297722
VERSION     AW297722.1  GI:6704358
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Dec 20, 1995 this sequence version replaced gi:1134505.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html The following repetitive
            elements were found in this cDNA sequence: 15-66, >MLT1E#LTR/MaLR
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2732001"
                     /clone_lib="NCI_CGAP_Sub6"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub6
                     is a subtracted library derived from BW, which consists of
                     a  mixture of four normalized libraries: NCI_CGAP_Brn50,
                     NCI_CGAP_Lu13,  NCI_CGAP_Ov18, GBC1 . The NCI_CGAP_Sub6
                     library had 7 million recombinants.  A single-stranded DNA
                     preparation of BW was used as a tracer in a subtractive
                     hybridization with a driver comprising: the IMAGE pool
                     (NCI_CGAP_Kid3 pool 1  LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775,1500552-1502855); NCI_CGAP_Kid5 pool 1
                     LLAM 3338-3342,3722-3725, 3776-3778 (IMAGE CloneIDs
                     1323912-1325831, 1471368-1472903, 1492104-1493255);
                     NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE
                     CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
                     pool 1 LLAM 3164-3167, 3716-3720, 3733-3735  (IMAGE
                     CloneIDs 1257096-1258631,1469064-1470983,
                     1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459,
                     2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
                     1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
                     LLAM 2644-2653,  2871-2872 (IMAGE CloneIDs
                     1057416-1061255,1144584-1145351). (50% of the driver
                     population), plus a pool of 3,840 arrayed clones from
                     NCI_CGAP_Sub1 (IMAGE  CloneIDs 2708616-2710535) and
                     NCI_CGAP_Sub2 (IMAGE CloneIDs 2710536-2712455)    (20% of
                     the driver population), plus a pool of 11,136 clones from
                     NCI_CGAP_Sub3  (IMAGE CloneIDs 2712456-2723591) (30% of
                     the driver population). Subtraction was performed as
                     previously described [Bonaldo, Lennon & Soares (1996):
                     Normalization and  Subtraction: Two Approaches To
                     Facilitate Gene Discovery. Genome Research 6, 791-806.
                     TAG_LIB=NCI_CGAP-Lu13
                     TAG_TISSUE=lung
                     TAG_SEQ=GCCGG"
BASE COUNT      109 a    111 c     88 g    160 t
ORIGIN


  Query Match             39.2%;  Score 29.4;  DB 79;  Length 468;
  Best Local Similarity   63.4%;  Pred. No. 3.2e+02;
  Matches   45;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
           || | |  |||||  ||| |  |||    |   ||| | ||| | |||| ||||| ||||
Db     162 GGTCCAGGATTGACTTTTCAGGTTTGGCTGGGGGACTTTAAGGAGGCCTCCATGAGCCCT 103

Qy      61 GCCGCCAACGG 71
           ||| ||   ||
Db     102 GCCTCCTGTGG 92


RESULT   9
AA339067/c
```

```
LOCUS       AA339067      230 bp    mRNA            EST       21-APR-1997
DEFINITION  EST44123 Fetal brain I Homo sapiens cDNA 5' end similar to EST
            containing Alu repeat, mRNA sequence.
ACCESSION   AA339067
VERSION     AA339067.1  GI:1991315
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 230)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., ,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   12140200
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393066.
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .230
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):140684"
                     /db_xref="taxon:9606"
                     /clone_lib="Fetal brain I"
                     /sex="female"
                     /dev_stage="fetus, 24 wks"
                     /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       52 a     67 c     48 g     62 t      1 others
ORIGIN


  Query Match             38.9%;  Score 29.2;  DB 32;  Length 230;
  Best Local Similarity   81.0%;  Pred. No. 4.7e+02;
  Matches   34;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAA 42
          |||||| |||| |||||  |||||  ||||||  ||||||||
Db    185 GGCCAACAATTAAAATTACATTTTAATTTTTTTAAATATAAA 144


RESULT  10
AI902252/c
LOCUS       AI902252      644 bp    mRNA            EST       01-DEC-1999
DEFINITION  QV-BT003-041198-029 BT003 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI902252
VERSION     AI902252.1  GI:6492730
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     On Dec 20, 1995 this sequence version replaced gi:1124918.
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT003-
            029.html&t3=041198&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT003"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      141 a    112 c    116 g    256 t     19 others
ORIGIN


  Query Match             38.9%;  Score 29.2;  DB 62;  Length 644;
  Best Local Similarity   62.2%;  Pred. No. 3.2e+02;
  Matches   46;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCG 60
          || ||||||   |  |||| ||||||||||||  || | ||| |  || ||     ||||
Db    270 GGGCAAAAAAAAAATCTTTTTTTTTTTTTTTTTTTAAAAAAAAACCCCCCCCCCCCCCCG 211

Qy     61 GCCGCCAACGGCGG 74
          | |     || ||
Db    210 GGGGGGGGGGGGGG 197


RESULT  11
CNS00F3L
LOCUS       CNS00F3L      1024 bp    DNA             GSS       04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR30J05 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069994
VERSION     AL069994.1  GI:4950125
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1024)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
```

```
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES              Location/Qualifiers
     source           1. .1024
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone_lib="RPCI-98"
                      /clone="BACR30J05"
                      /note="end : TET3"
BASE COUNT      260 a     172 c     189 g     339 t      64 others
ORIGIN


  Query Match             38.9%;  Score 29.2;  DB 82;  Length 1024;
  Best Local Similarity   43.8%;  Pred. No. 2.8e+02;
  Matches   32;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       2 GCCAAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGG 61
           ::||    ||: |||||| ||||||||||||    | :   :    ::| : :|: :||:
Db     898 KYCACCGNTTKNAATTTTTTTTTTTTTTTTTTTTTTTKKKKKKKNNNSSCGVGRGRRMCGS 957

Qy      62 CCGCCAACGGCGG 74
             : |::: || ||
Db     958 GSCCMRRGGGGGG 970


RESULT  12
AW026618
LOCUS       AW026618      368 bp    mRNA            EST       27-OCT-1999
DEFINITION  wv45c10.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2532498 3'
            similar to gb:X06409 RAF PROTO-ONCOGENE SERINE/THREONINE-PROTEIN
            KINASE (HUMAN);, mRNA sequence.
ACCESSION   AW026618
VERSION     AW026618.1  GI:5880237
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jul 7, 1999 this sequence version replaced gi:5405757.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40UP from Gibco
            High quality sequence stop: 208.
FEATURES              Location/Qualifiers
     source           1. .368
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2532498"
                      /clone_lib="NCI_CGAP_Gas4"
                      /tissue_type="poorly differentiated adenocarcinoma with
                      signet ring cell features"
                      /lab_host="DH10B"
                      /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.69 kb. Life Technologies catalog #:
                      11549-011"
BASE COUNT       95 a      70 c      65 g     135 t       3 others
ORIGIN


  Query Match             38.7%;  Score 29;  DB 64;  Length 368;
  Best Local Similarity   63.8%;  Pred. No. 4.4e+02;
  Matches   44;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCG 64
           ||| |  |   |||| || |||||||||  || | |||  ||  ||| || | ||  ||
Db     171 AAACAAGGCTGTTTTTTTGTTTTTTTTTTAAAAAAAAAAGGGTGGCCCTGCGGCCCCCCC 230

Qy      65 CCAACGGCG 73
             |||  || |
Db     231 CCATAGGGG 239


RESULT  13
AA123454/c
LOCUS       AA123454      378 bp    mRNA            EST       17-FEB-1997
DEFINITION  mp94a05.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:576848
            5', mRNA sequence.
ACCESSION   AA123454
VERSION     AA123454.1  GI:1681489
KEYWORDS    EST. .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:430248.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:351496
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 367.
FEATURES              Location/Qualifiers
     source           1. .378
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:576848"
                      /clone_lib="Soares_thymus_2NbMT"
                      /sex="male"
                      /tissue_type="Thymus"
                      /dev_stage="4 weeks"
```

```
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      116 a     91 c     81 g     90 t
ORIGIN

  Query Match             38.4%;  Score 28.8;  DB 28;  Length 378;
  Best Local Similarity   75.0%;  Pred. No. 4.9e+02;
  Matches   36;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      10 TTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGC 57
           || || | ||||||||||||||||  ||| ||||| ||   |  |||||
Db      64 TTTAATTCTTATTTTTTTTTTTTTAATTTAAAGAGGAACCTCTGGGC 17


RESULT  14
AI683617/c
LOCUS       AI683617      423 bp    mRNA            EST       15-DEC-1999
DEFINITION  tw52g05.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2263352 3'
            similar to contains element MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   AI683617
VERSION     AI683617.1  GI:4893799
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 423)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On May 18, 1998 this sequence version replaced gi:3138026.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1888   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 417.
FEATURES             Location/Qualifiers
     source          1. .423
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2263352"
                     /clone_lib="NCI_CGAP_Ut1"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
BASE COUNT      101 a    105 c     78 g    139 t
ORIGIN

  Query Match             38.4%;  Score 28.8;  DB 53;  Length 423;
  Best Local Similarity   65.6%;  Pred. No. 4.7e+02;
  Matches   42;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGCCCGGCCG 64
           |||||  | |||||| |  ||||||||||||      |||| |||| ||  |||   ||  |
Db     151 AAAAAGGGGAATTTTTTCCTTTTTTTTTGGGCCCCAAAGGGGCCCCCGGGGGTTGGGGGG 92

Qy      65 CCAA 68
             | |
Db      91 GCTA 88


RESULT  15
W30577/c
LOCUS       W30577        425 bp    mRNA            EST       11-SEP-1996
DEFINITION  mc28a08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:349814 5' similar to gb:U08378 Mus musculus acute phase
            response factor (MOUSE);, mRNA sequence.
ACCESSION   W30577
VERSION     W30577.1  GI:1310546
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:692920.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:221614
            Seq primer: ETPrimer
            High quality sequence stop: 423.
FEATURES             Location/Qualifiers
     source          1. .425
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:349814"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
BASE COUNT      125 a    106 c     91 g    103 t
ORIGIN

  Query Match             38.4%;  Score 28.8;  DB 26;  Length 425;
```

Best Local Similarity 75.0%; Pred. No. 4.7e+02;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy      10 TTGAAATTTTATTTTTTTTTTTTGGAATATAAAGCGGCCGCCATGGGC 57
           || || | |||||||||||||||  ||| ||||| ||    |  |||||
Db      91 TTTAATTCTTATTTTTTTTTTTTTTAATTTAAAGAGGAACCTCTGGGC 44
```

Search completed: May 29, 2000, 21:13:13
Job time: 36512 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:19 ; Search time 1214.87 Seconds
(without alignments)
909.139 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	70.8	48.8	4326	14	AF072997	AF072997 Cloning v
c	2	69	47.6	259	5	A44281	A44281 Sequence 9
	3	69	47.6	2958	14	ARBLKSM	X52326 pBluescript
	4	69	47.6	2958	14	ARBLKSP	X52331 pBluescript
	5	69	47.6	2961	14	ARBL2KSM	X52329 pBluescript
	6	69	47.6	2961	14	ARBL2KSP	X52327 pBluescript
	7	69	47.6	2964	14	SYNBLKSMV	L08784 BlueScribe
	8	69	47.6	2964	14	SYNBLKSPV	L08785 BlueScribe
c	9	69	47.6	3306	14	SYNPBEN66	D85525 Cloning vec
	10	69	47.6	4144	14	XXU35131	U35131 Plasmid pBS
	11	69	47.6	4267	14	PRS304	U03436 Yeast integ
	12	69	47.6	4289	14	XXU35136	U35136 Plasmid pBS
	13	69	47.6	4373	14	PRS306	U03438 Yeast integ
	14	69	47.6	4443	14	PRS303	U03435 Yeast integ
c	15	69	47.6	4670	14	ASAJ5326	AJ005326 pGAII(+)
	16	69	47.6	4670	14	ASAJ5329	AJ005329 pGAII(-)
	17	69	47.6	4707	14	XXU02374	U02374 Cloning vec
	18	69	47.6	4768	14	XXU25061	U25061 Cloning vec
	19	69	47.6	4783	14	PRS314	U03440 Yeast centr
	20	69	47.6	4887	14	PRS316	U03442 Yeast centr
	21	69	47.6	4950	14	XXU25060	U25060 Cloning vec
	22	69	47.6	4967	14	PRS313	U03439 Yeast centr
	23	69	47.6	5144	14	CVU23751	U23751 Cloning vec
	24	69	47.6	5187	14	U34887	U34887 Yeast integ
	25	69	47.6	5228	14	XXU25059	U25059 Cloning vec
	26	69	47.6	5504	14	PRS305	U03437 Yeast integ
	27	69	47.6	5634	14	CVU14125	U14125 Cloning vec
	28	69	47.6	6018	14	PRS315	U03441 Yeast centr
c	29	69	47.6	6340	14	ASAJ5323	AJ005323 pCP1(-) K
	30	69	47.6	9655	14	SYNPGR8V	AD001531 Cloning v
c	31	68.2	47.0	793	5	AR060142	AR060142 Sequence
c	32	67.8	46.8	338	1	BSPX91477	X91477 Bacterial s
	33	67.8	46.8	4323	14	AF072538	AF072538 Cloning v
c	34	67.4	46.5	3240	8	AF015771	AF015771 Magnaport
	35	67	46.2	144	14	SYNPLKRA	M22847 Cloning vec
c	36	67	46.2	147	14	SYNPLKRB	M22848 Cloning vec
c	37	67	46.2	799	4	AUU82676	U82676 Aphelocoma
c	38	67	46.2	2958	14	ARBLSKM	X52324 pBluescript
c	39	67	46.2	2958	14	ARBLSKP	X52325 pBluescript
c	40	67	46.2	2961	14	ARBL2SKM	X52330 pBluescript
c	41	67	46.2	2961	14	ARBL2SKP	X52328 pBluescript
c	42	67	46.2	2964	14	SYNBLDKPV	L08787 BlueScribe
c	43	67	46.2	2964	14	SYNBLSKMV	L08786 BlueScribe
c	44	67	46.2	3062	14	CVU47947	U47947 Cloning vec
c	45	67	46.2	3228	14	CVE7829	AJ007829 Cloning V

ALIGNMENTS

```
RESULT   1
AF072997
LOCUS       AF072997     4326 bp    DNA   circular  SYN       02-JUL-1998
DEFINITION  Cloning vector pCMVTAG3a, complete sequence.
ACCESSION   AF072997
VERSION     AF072997.1  GI:3284001
KEYWORDS    .
SOURCE      Cloning vector pCMVTAG3a.
  ORGANISM  Cloning vector pCMVTAG3a
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4326)
  AUTHORS   Lu,Q.
  TITLE     pCMVTAG3a
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4326)
  AUTHORS   Lu,Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-1998) Technical Services, Stratagene, 11011 N.
            Torrey Pines Rd., San Diego, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .4326
                     /organism="Cloning vector pCMVTAG3a"
                     /db_xref="taxon:78249"
                     /lab_host="Escherichia coli K12"
     promoter        1. .602
                     /note="CMV promoter"
     misc_feature    713. .799
                     /note="multiple cloning site"
     misc_feature    859. .1242
                     /note="SV40 polyA signal region"
     rep_origin      1242. .1703
                     /note="f1 origin of replication"
     promoter        1704. .1835
                     /note="Pbla promoter"
     rep_origin      1836. .2227
                     /note="SV40 origin of replication"
     CDS             2228. .3022
                     /function="kanamycin resistance"
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside phosphotransferase"
                     /protein_id="AAC25426.1"
                     /db_xref="GI:3284002"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKASMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
     rep_origin      3469. .4316
                     /note="ColE1 origin of replication"
BASE COUNT     1056 a   1116 c   1107 g   1047 t
ORIGIN
```

```
  Query Match             48.8%;  Score 70.8;  DB 14;  Length 4326;
  Best Local Similarity   76.3%;  Pred. No. 1.9e-12;
  Matches   87;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      12 GAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGCAGGAATTC 71
           | | ||    |||| |  |  | |   ||||   | ||||||||||||||||||||||||
Db     681 GCAGAAACTCATCTCTGAAGAGGATCTGGCCCGGGCGGATCCCCCGGGCTGCAGGAATTC 740

Qy      72 GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACTAACTAATT 125
           ||||||||||||||||||||||||||||||||||||||||        |||||
Db     741 GATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCTTAATT 794
```

```
RESULT   2
A44281/c
LOCUS       A44281       259 bp    DNA             PAT       07-MAR-1997
DEFINITION  Sequence 9 from Patent WO9511300.
ACCESSION   A44281
VERSION     A44281.1  GI:2299115
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Chandley,A.C., Kun,M., Sharkey,A.M., Hargreave,T.B. and Cooke,H.J.
  TITLE     AZOOSPERMIA IDENTIFICATION AND TREATMENT
  JOURNAL   Patent: WO 9511300-A 9 27-APR-1995;
            MEDICAL RES COUNCIL (GB)
COMMENT     Other publication AU 7947794 950508.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /clone_lib="MOUSE GENOMIC LIBRARY"
                     /clone="M3.2"
BASE COUNT       67 a     58 c     63 g     69 t      2 others
ORIGIN
```

```
  Query Match             47.6%;  Score 69;  DB 5;  Length 259;
  Best Local Similarity   88.2%;  Pred. No. 5.7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 53

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db      52 GGGGGCCCGGTACCCAGCTTTTGTT 28
```

```
RESULT   3
ARBLKSM
LOCUS       ARBLKSM      2958 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52326
VERSION     X52326.1  GI:58064
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(-)"
BASE COUNT      708 a    754 c    731 g    765 t
ORIGIN
```

```
  Query Match             47.6%;  Score 69;  DB 14;  Length 2958;
  Best Local Similarity   88.2%;  Pred. No. 7e-12;
```

```
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db    747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   4
ARBLKSP
LOCUS       ARBLKSP      2958 bp    DNA   circular  SYN       11-JUN-1998
DEFINITION  pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52331
VERSION     X52331.1  GI:58065
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
REFERENCE   4  (bases 1 to 2958)
  AUTHORS   Lampe,D.J., Grant,T.E. and Robertson,H.M.
  TITLE     Factors affecting transposition of the Himar1 mariner transposon in
            vitro
  JOURNAL   Genetics 149 (1), 179-187 (1998)
  MEDLINE   98250682
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(+)"
BASE COUNT      749 a    734 c    751 g    724 t
ORIGIN


  Query Match             47.6%;  Score 69;  DB 14;  Length 2958;
  Best Local Similarity   88.2%;  Pred. No. 7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db    747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   5
ARBL2KSM
LOCUS       ARBL2KSM     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(-) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52329
VERSION     X52329.1  GI:58060
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(-)"
BASE COUNT      706 a    758 c    735 g    762 t
ORIGIN


  Query Match             47.6%;  Score 69;  DB 14;  Length 2961;
  Best Local Similarity   88.2%;  Pred. No. 7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db    747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   6
ARBL2KSP
LOCUS       ARBL2KSP     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(+) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52327
VERSION     X52327.1  GI:58061
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
```

```
                 /db_xref="taxon:32630"
    misc_feature    1. .2961
                    /note="phagemid pBluescriptII KS(+)"
BASE COUNT      747 a    738 c    755 g    721 t
ORIGIN


  Query Match             47.6%;  Score 69;  DB 14;  Length 2961;
  Best Local Similarity   88.2%;  Pred. No. 7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   7
SYNBLKSMV
LOCUS       SYNBLKSMV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Minus cloning vector.
ACCESSION   L08784
VERSION     L08784.1  GI:310728
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Minus - Cloning vector
            ENTRY BLUEKSM                #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Minus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0077
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the minus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
               #parent
                 VecBase(3):BlueM13m
               #parent
                 VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                 VecSource(3):bGalKS, GenBank(50):PF1
               #brother
                 VecBase(3):BlueKSp, VecBase(3):BlueSKm
            PARENT
               Features of BlueKSm (2964 bp)
                 residue     source
                   3- 458  5488-5943     phage f1
                 460- 624   236- 400     pUC19c
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-2964   448-2617     pUC19c
               Conflict (cfl) and Mutations (mut): none
            PARENT
               Features of BlueKSm (2964 bp)
                 residue     source
                   3- 458  5488-5943     phage f1
                 460- 624   449- 285 (c) pUC19
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-1031   237-   1 (c) pUC19
                1032-2964  2686- 754 (c) pUC19
               Conflict (cfl) and Mutations (mut): none
            FEATURE
                     643           start of T7-RNA synthesis
                     774        (c) start of T3-RNA synthesis
               1976-2764 789-1 (c) Ap-R; b-lactamase
            POLYLINKER
                 KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
                 SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
               #resistance  Ap
               #indicator beta-galactosidase
            SUMMARY  BlueKSm    #length 2964   #checksum 1589.
FEATURES                Location/Qualifiers
     source             1. .2964
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
BASE COUNT      708 a    756 c    735 g    765 t
ORIGIN


  Query Match             47.6%;  Score 69;  DB 14;  Length 2964;
  Best Local Similarity   88.2%;  Pred. No. 7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   8
SYNBLKSPV
LOCUS       SYNBLKSPV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Plus cloning vector.
ACCESSION   L08785
VERSION     L08785.1  GI:310729
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Plus - Cloning vector
            ENTRY BLUEKSP                #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Plus - Cloning vector
            DATE    28-JAN-1987
```

```
      #sequence  02-FEB-1987
      #sequence  04-MAR-1987
      #sequence  03-APR-1987
   ACCESSION VB0078
   SOURCE    artificial
   REFERENCE
      #number 1
      #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
   J.
      #journal Gene (1987) in press
      #citation Sequence data from StrataGene
      #comment  sequence correction according to Stratagene COMMENT
      Obtained from StrataGene on floppy disc.
      Revised 02-FEB-1987 by F. Pfeiffer:
      1409/10 'AT' to 'TA' to match revised sequence of PBR322
      Revised 4-MAR-1987 to match sequence of pUC19 on request
      of Stratagene
      Polylinker region revised 03-APR-1987 according to Stratagene
   COMMENT
      The stand shown corresponds to pUC19c.
      As in the published sequence of pUC19c, The M13mp19 lacZ region
      is on the complementary strand.
   COMMENT
      This vector contains the f1 origin so that the plus strand
      can be obtained upon f1 superinfection.
   KEYWORDS
   CROSSREFERENCE
      #parent
        VecBase(3):BlueM13p
      #parent
        VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
        VecSource(3):bGalKS, GenBank(50):PF1
      #brother
        VecBase(3):BlueKSm, VecBase(3):BlueSKp
   PARENT
      Features of BlueKSp (2964 bp)
        residue     source
          3- 458  5943-5488 (c) phage f1
        460- 624   236- 400     pUC19c
        626- 645     1-  20     T7 promoter
        653- 760   108-   1 (c) BlueKS-polylinker
        772- 791    20-   1 (c) T3 promoter
        795-2964   448-2617     pUC19c
      Conflict (cfl) and Mutations (mut): none
   PARENT
      Features of BlueKSp (2964 bp)
        residue     source
          3- 458  5943-5488 (c) phage f1
        460- 624   449- 285 (c) pUC19
        626- 645     1-  20     T7 promoter
        653- 760   108-   1 (c) BlueKS-polylinker
        772- 791    20-   1 (c) T3 promoter
        795-1031   237-   1 (c) pUC19
       1032-2964  2686- 754 (c) pUC19
      Conflict (cfl) and Mutations (mut): none
   FEATURE
            643           start of T7-RNA synthesis
            774       (c) start of T3-RNA synthesis
      1976-2764 789-1 (c) Ap-R; b-lactamase
   POLYLINKER
        KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
        SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
      #resistance  Ap
      #indicator beta-galactosidase
   SUMMARY  BlueKSp    #length 2964    #checksum 690.
FEATURES             Location/Qualifiers
     source          1. .2964
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      750 a    736 c    755 g    723 t
ORIGIN
```

```
  Query Match            47.6%;  Score 69;  DB 14;  Length 2964;
  Best Local Similarity  88.2%;  Pred. No. 7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   9
SYNPBEN66/c
LOCUS       SYNPBEN66    3306 bp    DNA   circular  SYN       06-FEB-1999
DEFINITION  Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
            beta-lactamase, complete cds.
ACCESSION   D85525
VERSION     D85525.1  GI:1345433
KEYWORDS    plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE      Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
  ORGANISM  Cloning vector pBEN66
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 3306)
  AUTHORS   Yamamoto,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
            Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
            Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
            (Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE   2  (sites)
  AUTHORS   Yamamoto,Y. and Furuyama,J.
  TITLE     One-step disruption by circular DNA in Escherichia coli
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1. .3306
                     /organism="Cloning vector pBEN66"
                     /plasmid="pBEN66"
                     /db_xref="taxon:47800"
                     /lab_host="E.coli"
     promoter        19. .36
                     /note="T3 promoter"
     gene            260. .1075
                     /gene="kan"
     CDS             260. .1075
                     /gene="kan"
                     /note="derived from Tn903; kanamycin resistance gene"
                     /citation=[2]
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside 3'-phosphotransferase"
                     /protein_id="BAA12824.1"
                     /db_xref="GI:1345434"
                     /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                     GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                     GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                     VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
                     RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
     gene            complement(2209. .3069)
                     /gene="amp"
     CDS             complement(2209. .3069)
                     /gene="amp"
                     /note="amipicillin resistance gene"
                     /citation=[1]
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="BAA12825.1"
                     /db_xref="GI:1345435"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
```

DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
EIGASLIKHW"
BASE COUNT 854 a 800 c 790 g 862 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 3306;
Best Local Similarity 88.2%; Pred. No. 7e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 62

Qy 106 GGGGGCCTAACTAACTAATTTTGTT 130
||||||| | |||||||
Db 61 GGGGGCCCGGTACCCAGCTTTTGTT 37

RESULT 10
XXU35131
LOCUS XXU35131 4144 bp DNA circular SYN 26-SEP-1995
DEFINITION Plasmid pBSL159 cloning vector, complete sequence.
ACCESSION U35131
VERSION U35131.1 GI:984907
KEYWORDS .
SOURCE cloning vector pBSL159.
ORGANISM Cloning vector pBSL159
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4144)
AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
TITLE Improved antibiotic-resistance gene cassettes and omega elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis
JOURNAL Gene 160 (1), 63-67 (1995)
MEDLINE 95354958
REFERENCE 2 (bases 1 to 4144)
AUTHORS Hengen,P.N.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES Location/Qualifiers
source 1. .4144
/organism="Cloning vector pBSL159"
/plasmid="pBSL159"
/db_xref="taxon:42704"
CDS complement(1129. .1989)
/EC_number="3.5.2.6"
/codon_start=1
/transl_table=11
/function="ampicillin resistance"
/product="beta-lactamase"
/protein_id="AAC53618.1"
/db_xref="GI:984908"
/translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
EIGASLIKHW"
CDS complement(join(3704. .4144,1. .219))
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAC53619.1"
/db_xref="GI:984909"
/translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAF
LKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETF
SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNV
ANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGG
A"
BASE COUNT 1084 a 1009 c 959 g 1092 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4144;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2726 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2785

Qy 106 GGGGGCCTAACTAACTAATTTTGTT 130
||||||| | |||||||
Db 2786 GGGGGCCCGGTACCCAGCTTTTGTT 2810

RESULT 11
PRS304
LOCUS PRS304 4267 bp DNA circular SYN 14-SEP-1995
DEFINITION Yeast integrative vector pRS304 with TRP1 marker, complete sequence.
ACCESSION U03436
VERSION U03436.1 GI:416305
KEYWORDS .
SOURCE Cloning vector pRS304.
ORGANISM Cloning vector pRS304
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4267)
AUTHORS Sikorski,R.S. and Hieter,P.
TITLE A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910
REFERENCE 2 (bases 1 to 4267)
AUTHORS Stillman,D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral and Molecular Biology, University of Utah Medical Center, Salt Lake City, UT 84132, USA
FEATURES Location/Qualifiers
source 1. .4267
/organism="Cloning vector pRS304"
/db_xref="taxon:31827"
BASE COUNT 1128 a 965 c 1092 g 1082 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4267;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1980

Qy 106 GGGGGCCTAACTAACTAATTTTGTT 130
||||||| | |||||||
Db 1981 GGGGGCCCGGTACCCAGCTTTTGTT 2005

RESULT 12
XXU35136
LOCUS XXU35136 4289 bp DNA circular SYN 26-SEP-1995
DEFINITION Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION U35136
VERSION U35136.1 GI:984923
KEYWORDS .
SOURCE cloning vector pBSL97.
ORGANISM Cloning vector pBSL97
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4289)
AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
TITLE Improved antibiotic-resistance gene cassettes and omega elements

```
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4289)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4289
                     /organism="Cloning vector pBSL97"
                     /plasmid="pBSL97"
                     /db_xref="taxon:42709"
     CDS             complement(804. .1598)
                     /codon_start=1
                     /transl_table=11
                     /product="neomycin phosphotransferase"
                     /protein_id="AAC53629.1"
                     /db_xref="GI:984924"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
     CDS             complement(3301. .4161)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53630.1"
                     /db_xref="GI:984925"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     1024 a   1157 c   1111 g    997 t
ORIGIN


  Query Match           47.6%;  Score 69;  DB 14;  Length 4289;
  Best Local Similarity  88.2%;  Pred. No. 7.2e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2015 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2074

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db   2075 GGGGGCCCGGTACCCAGCTTTTGTT 2099


RESULT  13
PRS306
LOCUS       PRS306       4373 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS306 with URA3 marker, complete
            sequence.
ACCESSION   U03438
VERSION     U03438.1  GI:416307
KEYWORDS    .
SOURCE      Cloning vector pRS306.
  ORGANISM  Cloning vector pRS306
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4373)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4373)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4373
                     /organism="Cloning vector pRS306"
                     /db_xref="taxon:31829"
BASE COUNT     1181 a    987 c   1106 g   1099 t
ORIGIN


  Query Match           47.6%;  Score 69;  DB 14;  Length 4373;
  Best Local Similarity  88.2%;  Pred. No. 7.2e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2086

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db   2087 GGGGGCCCGGTACCCAGCTTTTGTT 2111


RESULT  14
PRS303
LOCUS       PRS303       4443 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS303 with HIS3 marker, complete
            sequence.
ACCESSION   U03435
VERSION     U03435.1  GI:416304
KEYWORDS    .
SOURCE      Cloning vector pRS303.
  ORGANISM  Cloning vector pRS303
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4443)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4443)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4443
                     /organism="Cloning vector pRS303"
                     /db_xref="taxon:31826"
BASE COUNT     1149 a   1048 c   1109 g   1137 t
ORIGIN


  Query Match           47.6%;  Score 69;  DB 14;  Length 4443;
  Best Local Similarity  88.2%;  Pred. No. 7.2e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2107 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db   2167 GGGGGCCCGGTACCCAGCTTTTGTT 2191


RESULT  15
ASAJ5326/c
```

```
LOCUS       ASAJ5326     4670 bp    DNA   circular  SYN       08-FEB-1999
DEFINITION  pGAII(+) KS positive selection cloning vector gltS gene.
ACCESSION   AJ005326
VERSION     AJ005326.1  GI:4028984
KEYWORDS    gltS gene; glutamate permease.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4670)
  AUTHORS   Gal,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
            Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
            6726, HUNGARY
REFERENCE   2  (bases 1 to 4670)
  AUTHORS   Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
            Kalman,M.
  TITLE     A positive selection cloning system based on the gltS gene of
            Escherichia coli
  JOURNAL   Anal. Biochem. 266 (2), 235-238 (1999)
  MEDLINE   99107575
FEATURES             Location/Qualifiers
     source          1. .4670
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     gene            852. .2231
                     /gene="gltS"
     CDS             852. .2231
                     /gene="gltS"
                     /codon_start=1
                     /transl_table=11
                     /product="glutamate permease"
                     /protein_id="CAA06473.1"
                     /db_xref="GI:4028985"
                     /translation="MFHLDTLATLVAATLTLLLGRKLVHSVSFLKKYTIPEPVAGGLL
                     VALALLVLKKSMGWEVNFDMSLRDPLMLAFFATIGLNANIASLRAGGRVVGIFLIVVV
                     GLLVMQNAIGIGMASLLGLDPLMGLLAGSITLSGGHGTGAAWSKLFIERYGFTNATEV
                     AMACATFGLVLGGLIGGPVARQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGDPLV
                     LERPPPRWSSNSPYSESYYARYLVKHSTTPNGIPDDQEVPTAFEKPDVGRMITSLVLI
                     ETIALIAICLTVGKIVAQLLAGTAFELPTFVCVLFVGVILSNGLSIMGFYRVFERAVS
                     VLGNVSLSLFLAMALMGLKLWELASLALPMLAILVVQTIFMALYAIFVTWRMMGKNYD
                     AAVLAAGHCGFGLGATPTAIANMQAITERFGPSHMAFLVVPMVGAFFIDIVNALVIKL
                     YLMLPIFAG"
BASE COUNT     1040 a   1165 c   1231 g   1234 t
ORIGIN

  Query Match           47.6%;  Score 69;  DB 14;  Length 4670;
  Best Local Similarity  88.2%;  Pred. No. 7.2e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1500 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1441

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db    1440 GGGGGCCCGGTACCCAGCTTTTGTT 1416


Search completed: May 29, 2000, 21:35:29
Job time: 36861 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:11 ; Search time 1446.77 Seconds
(without alignments)
25.075 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	69	47.6	259	1	Q87664	Mouse azoospermia
	2	69	47.6	501	1	T04866	Nucleotide analogu
	3	67	46.2	651	1	X20513	Polynucleotide seq
c	4	67	46.2	752	1	V31294	E. coli J96 pathog
c	5	67	46.2	1091	1	T75006	Human endogenous r
c	6	67	46.2	1122	1	T75005	Human endogenous r
c	7	67	46.2	3792	1	Q48463	Plasmid pG+host4 c
	8	67	46.2	4226	1	T39485	Human steroidogene
c	9	67	46.2	5234	1	Q48464	Plasmid pG+host5 c
c	10	67	46.2	6722	1	Q48465	Plasmid pG+host6 c
c	11	67	46.2	10529	1	V09028	Maize fluory2 gene
	12	67	46.2	12814	1	X24730	Swedish-FAD APP ta
	13	67	46.2	15692	1	X24731	London-FAD APP tar
	14	67	46.2	15692	1	X24732	Swedish/London-FAD
	15	67	46.2	15701	1	X24733	Swedish-FAD APP713
c	16	66.8	46.1	545	1	V68808	Human endogenous r
c	17	65.4	45.1	2640	1	Q26664	bDAT. cDNA encodin
c	18	65.4	45.1	5356	1	T43794	Plasmid pRIPHAT (r
	19	65	44.8	84	1	X02800	E. coli biotin DNA
c	20	65	44.8	3465	1	X02814	DE19731274 Seq ID
c	21	65	44.8	3481	1	X02815	DE19731274 Seq ID
c	22	65	44.8	3794	1	X02812	DE19731274 Seq ID
c	23	65	44.8	3810	1	X02813	DE19731274 Seq ID
	24	65	44.8	6596	1	V57377	Maize female-prefe
c	25	63.6	43.9	545	1	T75010	Human endogenous r
c	26	62.6	43.2	2150	1	T91037	Yeast checkpoint c
c	27	62.6	43.2	2150	1	X01271	Yeast RAD17 coding
	28	62.4	43.0	685	1	V10190	Stealth virus nucl
	29	62.4	43.0	685	1	V12003	Stealth virus plas
	30	60.2	41.5	1949	1	X05602	Nucleotide sequenc
	31	59	40.7	4164	1	T04575	Plasmid pAT-1 sequ
	32	59	40.7	4164	1	V22271	pAT-1 (pSD544). DN
	33	59	40.7	4164	1	V69740	Nucleotide sequenc
	34	58	40.0	5178	1	T49876	pTet-Splice. Nucle
	35	57.8	39.9	3198	1	T92702	Candida CaRho1 gen
	36	57.8	39.9	3198	1	T92869	Candida CaRho1 gen
	37	56.4	38.9	459	1	V57250	Clone #4 from muta
c	38	56	38.6	78	1	X02799	E. coli biotin DNA
	39	56	38.6	2973	1	V64254	Plasmid pKS varian
	40	56	38.6	3956	1	V64258	Plasmid pPK13/14 D
	41	56	38.6	4088	1	V64255	Plasmid pPK5/6 DNA
	42	56	38.6	4102	1	V64257	Plasmid pPK9/10 DN
	43	56	38.6	4583	1	V64256	Plasmid pPK7/8 DNA
	44	55	37.9	6505	1	T13394	Hybrid vector pSF-
	45	54	37.2	10504	1	V79503	DNA sequence of ex

ALIGNMENTS

RESULT 1
Q87664/c
ID Q87664 standard; DNA; 259 BP.
AC Q87664;
DT 06-MAR-1996 (first entry)
DE Mouse azoospermia factor (AZF) gene partial clone M3.2.
KW Azoospermia factor; AZF; male infertility; YRRM gene;
KW Y-chromosome; RNA recognition motif; ds.
OS Mus sp.
FH Key Location/Qualifiers
FT misc_difference 174
FT /*tag= a
FT /note= "base n at position 174 is not identified
FT in the specification"
FT misc_difference 197
FT /*tag= b
FT /note= "base n at position 197 is not identified
FT in the specification"
PN WO9511300-A2.
PD 27-APR-1995.
PF 24-OCT-1994; G02344.
PR 22-OCT-1993; GB-021857.
PR 07-JUL-1994; GB-013760.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;
DR WPI; 95-170221/22.
PT Nucleic acid encoding the human azoospermia factor, and probes and
PT antibodies specific for the sequence and encoded polypeptide - may
PT be used in the clinical diagnosis of male infertility
PS Disclosure; Fig 5; 40pp; English.
CC A mouse gene partial sequence (Q87664) showing homology to the human
CC azoospermia factor (AZF) YRMM gene was obtd. from mouse genomic phage
CC clone M3.2 isolated from a genomic library screened with MYK (YRRM1)
CC cDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm
CC of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also
CC obtd.
SQ Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;

Query Match 47.6%; Score 69; DB 1; Length 259;
Best Local Similarity 88.2%; Pred. No. 1.8e-15;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 53

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db      52 GGGGGCCCGGTACCCAGCTTTTGTT 28
```

RESULT 2
T04866
ID T04866 standard; DNA; 501 BP.
AC T04866;

```
DT   28-JAN-1996  (first entry)
DE   Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW   Nucleotide analogue; alkaline phosphatase; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .501
FT                   /*tag=  a
FT                   /note= "std. IUPAC codes used"
PN   WO9524505-A.
PD   14-SEP-1995.
PF   07-MAR-1995; G00490.
PR   08-MAR-1994; EP-301636.
PA   (AMSH ) AMERSHAM INT PLC.
PI   Reeve MA,  Robinson PS;
DR   WPI; 95-328290/42.
PT   Modification of residual fluorescence labelled nucleotide analogues  - to
PT   prevent migration in electrophoretic sequencing gel and interference with
PT   base calling of DNA chains
PS   Example; Fig 1; 18pp; English.
CC   The invention concerns the modification of redisual fluorescence
CC   labelled nt analogues to prevent migration in electrophoretic
CC   sequencing gel and interference with base calling of DNA chains. The
CC   modification involves the use of a phosphatase enzyme to remove at
CC   least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC   dye-terminating kt. The template used was 1 microg. of M13 mp8. The
CC   primer was M13 Universal primer. Samples were recovered. Some were
CC   subjected to calf intestinal alkaline phospatase digestion. This
CC   includes T04866. T04867 was a control nt. which was not treated with
CC   alkaline phosphatase. The example shows that alkaline phosphatase
CC   treatment causes removal of the dye-terminator artefacts and allows
CC   for accurate base calling with the ABI analysis software.
SQ   Sequence   501 BP;    98 A;     136 C;    140 G;    121 T;

  Query Match             47.6%;  Score 69;  DB 1;  Length 501;
  Best Local Similarity   88.2%;  Pred. No. 2.2e-15;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 116

Qy    106 GGGGGCCTAACTAACTAATTTTGTT 130
          |||||||       |   |||||||
Db    117 GGGGGCCCGGTACCCAGCTTTTGTT 141


RESULT   3
X20513
ID   X20513 standard; DNA; 651 BP.
AC   X20513;
DT   05-MAY-1999  (first entry)
DE   Polynucleotide sequence from the genome of Treponema pallidum.
KW   Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW   enzyme production; ds.
OS   Treponema pallidum.
PN   WO9859034-A2.
PD   30-DEC-1998.
PF   23-JUN-1998; U13041.
PR   24-JUN-1997; US-050667.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Fraser CM;
DR   WPI; 99-081273/07.
PT   New isolated Treponema pallidum nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of T. pallidum infections, particularly syphilis
PS   Claim 1; Page 257; 1150pp; English.
CC   X20500-21243 represent polynucleotide sequences from the genome of
CC   Treponema pallidum. The sequences can be used for detection,
CC   diagnosis, characterisation, prevention and therapy for T. pallidum
CC   infections, particularly syphilis. They can also be used for detecting
CC   diseases related to Borrelia infections in animals, and for the
CC   production of biosynthetic products such as enzymes.
SQ   Sequence   651 BP;    180 A;    167 C;    152 G;    149 T;

  Query Match             46.2%;  Score 67;  DB 1;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 137

Qy    106 GGGGGCC 112
          |||||||
Db    138 GGGGGCC 144


RESULT   4
V31294/c
ID   V31294 standard; DNA; 752 BP.
AC   V31294;
DT   01-OCT-1998  (first entry)
DE   E. coli J96 pathogenicity island contig #108.
KW   PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW   PAI V; pheV; vaccine; protective immune response; ds.
OS   Escherichia coli.
PN   WO9822575-A2.
PD   28-MAY-1998.
PF   21-NOV-1997; U21347.
PR   14-OCT-1997; US-061953.
PR   22-NOV-1996; US-031626.
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYWI-) UNIV WISCONSIN.
PI   Choi GH, Dillon PJ, Welch RA;
DR   WPI; 98-312461/27.
PT   New isolated uropathogenic E. coli nucleotide sequences - used to
PT   develop products for the detection of pathogenic E. coli and to
PT   elicit an immune response to pathogenic E. coli
PS   Claim 21; Page 221-222; 250pp; English.
CC   This sequence represents a E. coli strain J96 contig containing
CC   pathogenicity island (PAI) sequences, and represents a nucleic acid
CC   molecule of the invention. PAIs are large fragments of DNA which comprise
CC   pathogenicity determinants. The sequences of the invention are taken from
CC   PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC   on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC   approximately 94 min (at pheR) on the E. coli chromosome and is
CC   approximately 160 kb in size. Antibodies specific to the proteins encoded
CC   by the PAI open reading frames of the invention can be used in kits to
CC   detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC   a protective immune response in an animal to the uropathogenic E. coli
CC   strain J96.
SQ   Sequence   752 BP;    162 A;    213 C;    203 G;    172 T;

  Query Match             46.2%;  Score 67;  DB 1;  Length 752;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 13

Qy    106 GGGGGCC 112
          |||||||
Db     12 GGGGGCC 6


RESULT   5
T75006/c
ID   T75006 standard; DNA; 1091 BP.
AC   T75006;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 6.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
```

OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
DR WPI; 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 31-32; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (T75003-10) were obtd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1091 BP; 79 A; 350 C; 97 G; 248 T;

Query Match 46.2%; Score 67; DB 1; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 78

Qy 106 GGGGGCC 112
|||||||
Db 77 GGGGGCC 71

RESULT 6
T75005/c
ID T75005 standard; DNA; 1122 BP.
AC T75005;
DT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 3.
KW Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
DR WPI; 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 30-31; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (T75003-10) were obtd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1122 BP; 260 A; 316 C; 185 G; 279 T;

Query Match 46.2%; Score 67; DB 1; Length 1122;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 49

Qy 106 GGGGGCC 112
|||||||
Db 48 GGGGGCC 42

RESULT 7
Q48463/c
ID Q48463 standard; DNA; 3792 BP.
AC Q48463;
DT 18-MAR-1994 (first entry)
DE Plasmid pG+host4 containing Ts replication system.
KW Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
OS Synthetic.
PN WO9318164-A.
PD 16-SEP-1993.
PF 12-MAR-1993; F00248.
PR 13-MAR-1992; FR-003034.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Gruss A, Maguin E;
DR WPI; 93-303478/38.
PT New bacterial plasmid contg. heat sensitive replication system -
PT and marker gene, opt. capable of chromosomal integration, used to
PT inactivate specific gene or introduce heterologous gene
PS Example 2; Fig 9; 73pp; French.
CC Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC site, T7 and T3 promoters and sequencing primer binding sites. The
CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463). It
CC is thermosensitive in all hosts tested, including E.coli, and must be
CC maintained at 28 deg.C.
SQ Sequence 3792 BP; 1249 A; 552 C; 742 G; 1249 T;

Query Match 46.2%; Score 67; DB 1; Length 3792;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy 106 GGGGGCC 112
|||||||
Db 3481 GGGGGCC 3475

RESULT 8
T39485
ID T39485 standard; DNA; 4226 BP.
AC T39485;
DT 21-MAY-1997 (first entry)
DE Human steroidogenesis acute regulatory protein genomic DNA.
KW Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW mutation; detection; prenatal; genetic defect; congenital; protein;
KW lipoid adrenal hyperplasia; treatment; prevention; gene;
KW replacement therapy; hypercholesterolaemia; ds.
OS Homo sapiens.
PN WO9629338-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03896.
PR 23-MAR-1995; US-410540.
PA (REGC) UNIV CALIFORNIA.
PA (UYPE-) UNIV PENNSYLVANIA.
PI Lin D, Miller WL, Strauss JF;
DR WPI; 96-443130/44.
PT Isolated human steroidogenesis acute regulatory protein gene - used

```
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
PS   Claim 1; Pages 23-25; 89pp; English.
CC   The present sequence encodes the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc abd pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
SQ   Sequence   4226 BP;    940 A;    1132 C;    1144 G;    984 T;

  Query Match             46.2%;  Score 67;  DB 1;  Length 4226;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4149 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4208

Qy    106 GGGGGCC 112
          |||||||
Db   4209 GGGGGCC 4215


RESULT   9
Q48464/c
ID   Q48464 standard; DNA; 5234 BP.
AC   Q48464;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host5 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    36. .1496
FT                   /*tag=  a
FT                   /standard_name= ORI
FT                   /note= "origin of replication from pBR322"
FT   misc_feature    2640. .4383
FT                   /*tag=  b
FT                   /note= "from pGK12 (derived from pWV01)"
FT   misc_feature    4384. .4786
FT                   /*tag=  c
FT                   /note= "from pUB110"
FT   misc_feature    4787. .5234
FT                   /*tag=  d
FT                   /note= "from pSK"
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 10; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the 1.4kb AvaI-AlwNI fragment of
CC   pBR322 (containing the origin of replication) was inserted into
CC   NsiI-cleaved pG+host4 to give pG+host5 (Q48464).The pBR322 ORI
CC   allows maintenance of the plasmid in E.coli at 37 deg.C. while the
CC   heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
CC   bacteria.
SQ   Sequence   5234 BP;    1569 A;    973 C;    1133 G;    1559 T;


  Query Match             46.2%;  Score 67;  DB 1;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy    106 GGGGGCC 112
          |||||||
Db   4923 GGGGGCC 4917


RESULT  10
Q48465/c
ID   Q48465 standard; DNA; 6722 BP.
AC   Q48465;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host6 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Disclosure; Fig 11; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the AvaI-EcoRI fragment of pBR322
CC   (containing the origin of replication and the ampicillin resistance
CC   gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
CC   pG+host6 (Q48466).The pBR322 ORI allows maintenance of the plasmid in
CC   E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC   at 28 deg.C in gram-positive bacteria.
SQ   Sequence   6722 BP;    1961 A;    1278 C;    1503 G;    1980 T;


  Query Match             46.2%;  Score 67;  DB 1;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy    106 GGGGGCC 112
          |||||||
Db   6411 GGGGGCC 6405
```

```
RESULT  11
V09028/c
ID   V09028 standard; DNA; 10529 BP.
AC   V09028;
DT   20-JUL-1998  (first entry)
DE   Maize fluory2 gene (fl2).
KW   Floury2; fl2 gene; alpha-zein; signal peptide; transgenic plant;
KW   seed; essential amino acid; animal feed; maize; rice; wheat;
KW   barley; millet; sorghum; ds.
OS   Zea mays.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .760
FT                   /*tag= a
FT                   /note= "vector-derived sequence"
FT   promoter        761. .3824
FT                   /*tag= b
FT                   /note= "Claim 14"
FT   CDS             3825. .4613
FT                   /*tag= c
FT   sig_peptide     3825. .3887
FT                   /*tag= d
FT   mat_peptide     3888. .4610
FT                   /*tag= e
PN   WO9802563-A1.
PD   22-JAN-1998.
PF   11-JUL-1997; U11723.
PR   17-JUL-1996; US-021833.
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (UYAR-) UNIV ARIZONA STATE.
PI   Beach L, Coleman CE, Larkins BA;
DR   WPI; 98-110609/10.
DR   P-PSDB; W23977.
PT   Cereal plants containing trans-gene expressing fusion that includes
PT   signal peptide of the fl2 maize gene - and protein having high
PT   content of essential amino acids, producing feeds of improved
PT   nutritional value
PS   Claim 14; Fig 1A-H; 37pp; English.
CC   This is the nucleotide sequence of a clone of the fluory2 (fl2)
CC   gene of maize.  It codes for a 24-kDa alpha-zein protein (see
CC   W23977) that includes a 21-amino acid signal peptide (see W23976)
CC   which targets the alpha-zein to the lumen of the rough endoplasmic
CC   reticulum.  A claimed cereal plant contains a transgene comprising
CC   a first polynucleotide that encodes the fl2 signal peptide and a
CC   second polynucleotide that encodes an agronomically high-value
CC   protein.  Also new are seeds produced by the plants, the transgene
CC   itself, and a transgene that also includes the fl2 promoter.  The
CC   second polynucleotide preferably encodes a protein that has a high
CC   content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC   maize, wheat, rice, barley, millet or sorghum will have increased
CC   contents of these essential amino acids in their seeds (all
CC   claimed).
SQ   Sequence   10529 BP;    2845 A;    2484 C;    2262 G;    2938 T;


  Query Match             46.2%;  Score 67;  DB 1;  Length 10529;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 667

Qy    106 GGGGGCC 112
          |||||||
Db    666 GGGGGCC 660


RESULT  12
X24730
ID   X24730 standard; DNA; 12814 BP.
AC   X24730;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP targetting vector pMTI-2398.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2398;
KW   Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1932. .2276
FT                   /*tag= a
FT                   /standard_name= Swedish-FAD APP
FT                   /note= "encodes W97997"
FT   mat_peptide     5369. .6160
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97997, W97998.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 96-103; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP targetting
CC   vector pMTI-2398.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5007,
CC   created using vector pMTI-2398, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced.  The targetted Swedish-FAD
CC   m/hAPP gene expressed m/hAPP protein at levels approaching those
CC   observed for mouse APP in brain.
SQ   Sequence   12814 BP;    1865 A;    1850 C;    1962 G;    1939 T;


  Query Match             46.2%;  Score 67;  DB 1;  Length 12814;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9888 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 9947

Qy    106 GGGGGCC 112
          |||||||
Db   9948 GGGGGCC 9954


RESULT  13
X24731
ID   X24731 standard; DNA; 15692 BP.
AC   X24731;
DT   21-JUN-1999  (first entry)
DE   London-FAD APP targetting vector pMTI-2453.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

```
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2453;
KW   London-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /standard_name= "Swedish-FAD"
FT                   /note= "encodes W97999"
FT   mutation        replace(4990,"")
FT                   /*tag= b
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= c
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W97999.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 104-113; 209pp; English.
CC   This is the nucleotide sequence of London-FAD APP targetting
CC   vector pMTI-2453. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5401,
CC   created using vector pMTI-2453, the London mutation. i.e.
CC   V(717)L, was also introduced.
SQ   Sequence   15692 BP;    2423 A;    2730 C;    2786 G;    2541 T;

  Query Match             46.2%;  Score 67;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12766 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12825

Qy    106 GGGGGCC 112
          |||||||
Db  12826 GGGGGCC 12832


RESULT  14
X24732
ID   X24732 standard; DNA; 15692 BP.
AC   X24732;
DT   21-JUN-1999  (first entry)
DE   Swedish/London-FAD APP targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2454;
KW   London-FAD APP; Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /note= "encodes W98000"
FT   mutation        replace(4849,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4989,"")
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98000.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 114-123; 209pp; English.
CC   This is the nucleotide sequence of Swedish/London-FAD APP targetting
CC   vector pMTI-2454. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5103,
CC   created using vector pMTI-2454, the London mutation. i.e.
CC   V(717)L, and the Swedish mutation, i.e. KM(670,671)NL, were
CC   introduced. The targetted Swedish/London-FAD m/hAPP gene expressed
CC   m/hAPP protein at levels approaching those observed for mouse APP
CC   in brain.
SQ   Sequence   15692 BP;    2423 A;    2731 C;    2783 G;    2542 T;

  Query Match             46.2%;  Score 67;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12766 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12825

Qy    106 GGGGGCC 112
          |||||||
Db  12826 GGGGGCC 12832


RESULT  15
X24733
ID   X24733 standard; DNA; 15701 BP.
AC   X24733;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP713 targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2455;
KW Swedish-FAD APP713; mutation; ss.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT mat_peptide 4807. .4983
FT /*tag= a
FT /note= "encodes W98001"
FT mutation replace(4835,"")
FT /*tag= b
FT /standard_name= "Swedish-FAD"
FT mutation replace(4981,"")
FT /*tag= c
FT /standard_name= APP713stop
FT mat_peptide 8232. .9032
FT /*tag= d
FT /note= "encodes W97998"
PN WO9909150-A1.
PD 25-FEB-1999.
PF 18-AUG-1997; U14507.
PR 18-AUG-1997; WO-U14507.
PA (FARB) BAYER CORP.
PI Wirak DO;
DR WPI; 99-181029/15.
DR P-PSDB; W97998, W98001.
PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease
PS Example; Page 124-133; 209pp; English.
CC This is the nucleotide sequence of Swedish-FAD APP713 targetting
CC vector pMTI-2455. The invention provides a novel gene targetting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).
CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targetting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human cDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5215,
CC created using vector pMTI-2455, the Swedish mutation. i.e.
CC KM(670,681)NL, was also introduced, and also a stop codon (T to
CC stop at position 714).
SQ Sequence 15701 BP; 2470 A; 2675 C; 2845 G; 2497 T;

Query Match 46.2%; Score 67; DB 1; Length 15701;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12775 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12834

Qy     106 GGGGGCC 112
           |||||||
Db   12835 GGGGGCC 12841
```

Search completed: May 29, 2000, 21:58:21
Job time: 38169 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:08:54 ; Search time 621.83 Seconds
(without alignments)
30.310 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	68.2	47.0	793	3	US-08-967-101-121	Sequence 121, App
c	2	68.2	47.0	793	4	US-08-592-541-121	Sequence 121, App
c	3	67	46.2	3792	4	US-08-992-334-1	Sequence 1, Appli
c	4	67	46.2	3792	5	US-08-302-752-1	Sequence 1, Appli
	5	67	46.2	4016	2	US-08-410-540-3	Sequence 3, Appli
c	6	67	46.2	5234	4	US-08-992-334-2	Sequence 2, Appli
c	7	67	46.2	5234	5	US-08-302-752-2	Sequence 2, Appli
c	8	67	46.2	6722	4	US-08-992-334-3	Sequence 3, Appli
c	9	67	46.2	6722	5	US-08-302-752-3	Sequence 3, Appli
	10	65.4	45.1	9318	3	US-08-793-610-6	Sequence 6, Appli
c	11	62.6	43.2	1023	1	US-08-198-446B-1	Sequence 1, Appli
c	12	62.6	43.2	1023	3	US-08-870-693-1	Sequence 1, Appli
c	13	62.6	43.2	2150	1	US-08-198-446B-10	Sequence 10, Appl
c	14	62.6	43.2	2150	3	US-08-870-693-10	Sequence 10, Appl
	15	62.4	43.0	685	2	US-08-463-115-56	Sequence 56, Appl
	16	62.4	43.0	685	2	US-08-465-388-56	Sequence 56, Appl
	17	59	40.7	4164	1	US-08-204-675-1	Sequence 1, Appli
	18	59	40.7	4164	3	US-08-660-754-1	Sequence 1, Appli
	19	59	40.7	4164	4	US-08-796-364-1	Sequence 1, Appli
	20	59	40.7	4164	6	PCT-US95-02520-1	Sequence 1, Appli
	21	58	40.0	5178	3	US-08-474-169-2	Sequence 2, Appli
	22	56	38.6	88	1	US-08-144-602B-15	Sequence 15, Appl
c	23	55	37.9	834	3	US-08-967-101-113	Sequence 113, App
c	24	55	37.9	834	4	US-08-592-541-113	Sequence 113, App
	25	55	37.9	6505	3	US-08-793-610-5	Sequence 5, Appli
	26	54	37.2	5534	2	US-08-452-267-3	Sequence 3, Appli
	27	50.6	34.9	3341	3	US-08-868-577-18	Sequence 18, Appl
c	28	50	34.5	780	3	US-08-967-101-116	Sequence 116, App
c	29	50	34.5	780	4	US-08-592-541-116	Sequence 116, App
c	30	50	34.5	2308	1	US-08-325-071-62	Sequence 62, Appl
c	31	49.8	34.3	1200	1	US-08-011-398B-3	Sequence 3, Appli
c	32	49.8	34.3	1200	2	US-08-464-051-3	Sequence 3, Appli
c	33	49.8	34.3	1200	3	US-08-462-498-3	Sequence 3, Appli
c	34	49.8	34.3	6244	1	US-08-076-726-15	Sequence 15, Appl
c	35	49.8	34.3	6244	1	US-08-260-452-8	Sequence 8, Appli
c	36	49.8	34.3	6244	3	US-08-481-970-8	Sequence 8, Appli
c	37	49.8	34.3	6244	4	US-08-897-719-8	Sequence 8, Appli
	38	48	33.1	1612	1	US-08-343-733A-1	Sequence 1, Appli
	39	48	33.1	2764	4	US-08-465-971B-1	Sequence 1, Appli
c	40	47.2	32.6	2185	4	US-08-467-948A-3	Sequence 3, Appli
c	41	47	32.4	1997	4	US-08-667-809B-3	Sequence 3, Appli
c	42	47	32.4	2126	3	US-08-789-354-1	Sequence 1, Appli
	43	47	32.4	2605	3	US-08-680-395-4	Sequence 4, Appli
c	44	46.6	32.1	1249	2	US-08-463-115-35	Sequence 35, Appl
c	45	46.6	32.1	1249	2	US-08-465-388-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-967-101-121

Query Match 47.0%; Score 68.2; DB 3; Length 793;
Best Local Similarity 85.4%; Pred. No. 2.7e-14;
Matches 76; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||      |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   2
US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-592-541-121


  Query Match             47.0%;  Score 68.2;  DB 4;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 2.7e-14;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           .  ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||      |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   3
US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
;    IMMEDIATE SOURCE:
;      CLONE:  pG+host4
US-08-992-334-1


  Query Match             46.2%;  Score 67;  DB 4;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy     106 GGGGGCC 112
           |||||||
Db    3481 GGGGGCC 3475


RESULT   4
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-1


  Query Match             46.2%;  Score 67;  DB 5;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy     106 GGGGGCC 112
           |||||||
Db    3481 GGGGGCC 3475


RESULT   5
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F.
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,540
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-238/00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4016 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1098..1283
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1620..1733
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2047..2174
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2267..2425
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2567..2751
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2828..2921
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3031..3765
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1433..1434
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2208..2209
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2781..2782
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2956..2957
;      OTHER INFORMATION:  /note= "interruption of sequence data"
US-08-410-540-3


  Query Match             46.2%;  Score 67;  DB 2;  Length 4016;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3939 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3998

Qy     106 GGGGGCC 112
           |||||||
Db    3999 GGGGGCC 4005


RESULT   6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
```

```
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-2


  Query Match             46.2%;  Score 67;  DB 4;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy     106 GGGGGCC 112
           |||||||
Db    4923 GGGGGCC 4917


RESULT   7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-2


  Query Match             46.2%;  Score 67;  DB 5;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy     106 GGGGGCC 112
           |||||||
Db    4923 GGGGGCC 4917


RESULT   8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-3
```

```
  Query Match             46.2%;  Score 67;  DB 4;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy     106 GGGGGCC 112
           |||||||
Db    6411 GGGGGCC 6405


RESULT   9
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-3


  Query Match             46.2%;  Score 67;  DB 5;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy     106 GGGGGCC 112
           |||||||
Db    6411 GGGGGCC 6405


RESULT  10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6


  Query Match             45.1%;  Score 65.4;  DB 3;  Length 9318;
  Best Local Similarity   98.5%;  Pred. No. 4.9e-13;
  Matches   66;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5654 GGGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 5713

Qy     106 GGGGGCC 112
           |||||||
Db    5714 GGGGGCC 5720


RESULT  11
US-08-198-446B-1/c
; Sequence 1, Application US/08198446B
; Patent No. 5674996
;  GENERAL INFORMATION:
;    APPLICANT:  Hartwell, Leland H.
;    APPLICANT:  Weinert, Ted A.
;    APPLICANT:  Plon, Sharon E.
;    APPLICANT:  Groudine, Mark T.
;    TITLE OF INVENTION:  Cell Cycle Checkpoint Genes
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen O'Connor Johnson Kindness PLLC
;      STREET:  1420 Fifth Ave., Suite 2800
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/198,446B
;      FILING DATE:  18-FEB-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sheiness, Diana K.
;      REGISTRATION NUMBER:  35,356
;      REFERENCE/DOCKET NUMBER:  FHCR17537
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-682-8100
;      TELEFAX:  206-224-0779
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1023 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  yeast RAD17 cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
US-08-198-446B-1


  Query Match             43.2%;  Score 62.6;  DB 1;  Length 1023;
  Best Local Similarity   73.4%;  Pred. No. 2.2e-12;
  Matches   80;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGC 63
          | |||||   |||| | |    ||||  | ||   |  |     | ||     |||||||
Db    204 CGAAAATGATAAAAATTGGGAAATTTGATCCATTTGAACAAATGGCATTTTAGGGGCTGC 145

Qy     64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 112
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 96


RESULT  12
US-08-870-693-1/c
; Sequence 1, Application US/08870693
; Patent No. 5866338
;  GENERAL INFORMATION:
;    APPLICANT:  Hartwell, Leland H.
;    APPLICANT:  Weinert, Ted A.
;    APPLICANT:  Plon, Sharon E.
;    APPLICANT:  Groudine, Mark T.
;    TITLE OF INVENTION:  Cell Cycle Checkpoint Genes
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen O'Connor Johnson Kindness PLLC
;      STREET:  1420 Fifth Ave., Suite 2800
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/870,693
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/198,446
;      FILING DATE:  February 18, 1994
;      APPLICATION NUMBER:  PCT/US93/04458
;      FILING DATE:  May 12, 1993
;      APPLICATION NUMBER:  US 07/884,426
;      FILING DATE:  May 14, 1992
;      APPLICATION NUMBER:  US 07/882,051
;      FILING DATE:  May 12, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sheiness, Diana K.
;      REGISTRATION NUMBER:  35,356
;      REFERENCE/DOCKET NUMBER:  FHCR110798
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-682-8100
;      TELEFAX:  206-224-0779
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1023 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  yeast RAD17 cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
US-08-870-693-1


  Query Match             43.2%;  Score 62.6;  DB 3;  Length 1023;
  Best Local Similarity   73.4%;  Pred. No. 2.2e-12;
  Matches   80;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGC 63
          | |||||   |||| | |    ||||  | ||   |  |     | ||     |||||||
Db    204 CGAAAATGATAAAAATTGGGAAATTTGATCCATTTGAACAAATGGCATTTTAGGGGCTGC 145

Qy     64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 112
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 96


RESULT  13
US-08-198-446B-10/c
; Sequence 10, Application US/08198446B
; Patent No. 5674996
;  GENERAL INFORMATION:
;    APPLICANT:  Hartwell, Leland H.
;    APPLICANT:  Weinert, Ted A.
;    APPLICANT:  Plon, Sharon E.
;    APPLICANT:  Groudine, Mark T.
;    TITLE OF INVENTION:  Cell Cycle Checkpoint Genes
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen O'Connor Johnson Kindness PLLC
;      STREET:  1420 Fifth Ave., Suite 2800
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/198,446B
;      FILING DATE:  18-FEB-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sheiness, Diana K.
;      REGISTRATION NUMBER:  35,356
;      REFERENCE/DOCKET NUMBER:  FHCR17537
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-682-8100
;      TELEFAX:  206-224-0779
```

```
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  yeast RAD17 cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  741..1947
US-08-198-446B-10


  Query Match             43.2%;  Score 62.6;  DB 1;  Length 2150;
  Best Local Similarity   73.4%;  Pred. No. 2.8e-12;
  Matches   80;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGC 63
          | |||||   |||| | |    ||||  | ||   |  |     | ||     |||||||
Db    204 CGAAAATGATAAAAATTGGGAAATTTGATCCATTTGAACAAATGGCATTTTAGGGGCTGC 145

Qy     64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 112
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 96


RESULT  14
US-08-870-693-10/c
; Sequence 10, Application US/08870693
; Patent No. 5866338
;  GENERAL INFORMATION:
;    APPLICANT:  Hartwell, Leland H.
;    APPLICANT:  Weinert, Ted A.
;    APPLICANT:  Plon, Sharon E.
;    APPLICANT:  Groudine, Mark T.
;    TITLE OF INVENTION:  Cell Cycle Checkpoint Genes
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen O'Connor Johnson Kindness PLLC
;      STREET:  1420 Fifth Ave., Suite 2800
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/870,693
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/198,446
;      FILING DATE:  February 18, 1994
;      APPLICATION NUMBER:  PCT/US93/04458
;      FILING DATE:  May 12, 1993
;      APPLICATION NUMBER:  US 07/884,426
;      FILING DATE:  May 14, 1992
;      APPLICATION NUMBER:  US 07/882,051
;      FILING DATE:  May 12, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sheiness, Diana K.
;      REGISTRATION NUMBER:  35,356
;      REFERENCE/DOCKET NUMBER:  FHCR110798
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-682-8100
;      TELEFAX:  206-224-0779
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  yeast RAD17 cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  741..1947
US-08-870-693-10


  Query Match             43.2%;  Score 62.6;  DB 3;  Length 2150;
  Best Local Similarity   73.4%;  Pred. No. 2.8e-12;
  Matches   80;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      4 CAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGC 63
          | |||||   |||| | |    ||||  | ||   |  |     | ||     |||||||
Db    204 CGAAAATGATAAAAATTGGGAAATTTGATCCATTTGAACAAATGGCATTTTAGGGGCTGC 145

Qy     64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 112
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 96


RESULT  15
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READAELE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,115
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:            Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5703221ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
```

; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
US-08-463-115-56

Query Match 43.0%; Score 62.4; DB 2; Length 685;
Best Local Similarity 94.0%; Pred. No. 2.3e-12;
Matches 63; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||||
Db     150 GTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGAGG 209

Qy     106 GGGGGCC 112
           ||||| |
Db     210 GGGGGNC 216
```

Search completed: May 29, 2000, 22:09:00
Job time: 38741 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:27 ; Search time 2276.24 Seconds
(without alignments)
284.222 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145	100.0	145	28	US-08-935-377-6	Sequence 6, Appli
2	132	91.0	148	28	US-08-935-377-7	Sequence 7, Appli

	3	131	90.3	149	28	US-08-935-377-8	Sequence 8, Appli
	4	130	89.7	150	28	US-08-935-377-9	Sequence 9, Appli
	5	69	47.6	224	13	US-08-377-383-4	Sequence 4, Appli
	6	69	47.6	224	18	US-08-693-573-4	Sequence 4, Appli
	7	68.6	47.3	858	40	US-09-123-912-91	Sequence 91, Appl
	8	68.6	47.3	858	43	US-09-221-107-91	Sequence 91, Appl
	9	68.6	47.3	858	46	US-09-285-479-91	Sequence 91, Appl
c	10	68.2	47.0	793	15	US-08-431-048A-121	Sequence 121, App
c	11	68.2	47.0	793	16	US-08-496-841-121	Sequence 121, App
c	12	68.2	47.0	793	17	US-08-509-359-121	Sequence 121, App
c	13	68.2	47.0	793	17	US-08-509-359B-121	Sequence 121, App
c	14	68.2	47.0	793	40	US-09-124-523-121	Sequence 121, App
c	15	68.2	47.0	793	40	US-09-124-698-121	Sequence 121, App
c	16	68.2	47.0	793	40	US-09-127-480-121	Sequence 121, App
c	17	67	46.2	147	38	US-09-072-433-31	Sequence 31, Appl
c	18	67	46.2	157	63	US-60-029-960-595	Sequence 595, App
	19	67	46.2	245	54	US-09-411-999-19130	Sequence 19130, A
	20	67	46.2	381	49	US-09-332-782-8570	Sequence 8570, Ap
	21	67	46.2	381	57	US-09-515-694-8570	Sequence 8570, Ap
	22	67	46.2	393	54	US-09-411-999-13627	Sequence 13627, A
	23	67	46.2	405	49	US-09-332-782-8523	Sequence 8523, Ap
	24	67	46.2	405	57	US-09-515-694-8523	Sequence 8523, Ap
	25	67	46.2	420	49	US-09-332-782-8595	Sequence 8595, Ap
	26	67	46.2	420	57	US-09-515-694-8595	Sequence 8595, Ap
	27	67	46.2	422	49	US-09-332-782-9309	Sequence 9309, Ap
	28	67	46.2	422	57	US-09-515-694-9309	Sequence 9309, Ap
	29	67	46.2	423	54	US-09-411-999-34199	Sequence 34199, A
	30	67	46.2	433	42	US-09-179-473-2334	Sequence 2334, Ap
	31	67	46.2	433	49	US-09-328-351-2334	Sequence 2334, Ap
	32	67	46.2	442	42	US-09-179-473-2495	Sequence 2495, Ap
	33	67	46.2	442	49	US-09-328-351-2495	Sequence 2495, Ap
	34	67	46.2	448	42	US-09-179-473-2255	Sequence 2255, Ap
	35	67	46.2	448	49	US-09-328-351-2255	Sequence 2255, Ap
	36	67	46.2	454	42	US-09-179-473-2100	Sequence 2100, Ap
	37	67	46.2	454	42	US-09-179-473-2360	Sequence 2360, Ap
	38	67	46.2	454	49	US-09-328-351-2100	Sequence 2100, Ap
	39	67	46.2	454	49	US-09-328-351-2360	Sequence 2360, Ap
c	40	67	46.2	752	32	US-08-976-259-108	Sequence 108, App
c	41	67	46.2	752	64	US-60-031-626-108	Sequence 108, App
c	42	67	46.2	827	82	US-60-126-265-996	Sequence 996, App
c	43	67	46.2	2958	37	US-09-040-961-6	Sequence 6, Appli
c	44	67	46.2	2958	41	US-09-130-115-6	Sequence 6, Appli
c	45	67	46.2	2961	15	US-08-446-935-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT   1
US-08-935-377-6
; Sequence 6, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-6


  Query Match           100.0%;  Score 145;  DB 28;  Length 145;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-38;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGC 60

Qy      61 TGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACTAAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACTAAC 120

Qy     121 TAATTTTGTTTTTGTGGGCCCGGCC 145
           |||||||||||||||||||||||||
Db     121 TAATTTTGTTTTTGTGGGCCCGGCC 145


RESULT   2
US-08-935-377-7
; Sequence 7, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  148 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-7
```

```
  Query Match             91.0%;  Score 132;  DB 28;  Length 148;
  Best Local Similarity   98.0%;  Pred. No. 2.9e-34;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCC---GTGGATCCCCCG 57
          ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCG 60

Qy     58 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 120

Qy    118 AACTAATTTTGTTTTTGTGGGCCCGGCC 145
          ||||||||||||||||||||||||||||
Db    121 AACTAATTTTGTTTTTGTGGGCCCGGCC 148


RESULT   3
US-08-935-377-8
; Sequence 8, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-8


  Query Match             90.3%;  Score 131;  DB 28;  Length 149;
  Best Local Similarity   97.3%;  Pred. No. 6.3e-34;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCC----GTGGATCCCCC 56
          ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60

Qy     57 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120

Qy    117 TAACTAATTTTGTTTTTGTGGGCCCGGCC 145
          ||||||||||||||||||||||||||||
Db    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149


RESULT   4
US-08-935-377-9
; Sequence 9, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-9


  Query Match             89.7%;  Score 130;  DB 28;  Length 150;
  Best Local Similarity   96.7%;  Pred. No. 1.4e-33;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGC-----CGTGGATCCCC 55
          |||||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60

Qy     56 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120

Qy    116 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 145
          ||||||||||||||||||||||||||||
Db    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150


RESULT   5
US-08-377-383-4
; Sequence 4, Application US/08377383
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru, FUKAMI,Tadashi, HIYOSHI,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of
;    TITLE OF INVENTION:  Nucleic Acid
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman Darby & Cushman
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
```

```
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3918
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskett, 3.5 inch, 720Kb storage
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1 or ASCII editors
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/377,383
;       FILING DATE:  25-JAN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP6971/1994
;       FILING DATE:  26-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Glenn J. Perry
;       REGISTRATION NUMBER:  28458
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)861-3000
;       TELEFAX:  (202)822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  224 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-377-383-4


  Query Match             47.6%;  Score 69;  DB 13;  Length 224;
  Best Local Similarity   88.2%;  Pred. No. 4.2e-13;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db     148 GGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT   6
US-08-693-573-4
; Sequence 4, Application US/08693573
;  GENERAL INFORMATION:
;     APPLICANT:  HOSOI,Shigeru,
;     APPLICANT:  FUKAMI,Tadashi,
;     APPLICANT:  KOJIMA,Makiko
;     TITLE OF INVENTION:  Method of Determining Base Sequence of                Nucleic
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CUSHMAN DARBY CUSHMAN
;       ADDRESSEE:  INTELLECTUAL PROPERTY GROUP OF
;       ADDRESSEE:  PILLSBURY MADISON & SUTRO, L.L.P.
;       STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3918
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.5 inch, 720Kb storage
;       COMPUTER:  IBM PC/XT/AT compatibles
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1 or ASCII editors
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/693,573
;       FILING DATE:  07-AUG-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/377,383
;       FILING DATE:  25-JAN-1995
;       APPLICATION NUMBER:  JP6971/1994
;       FILING DATE:  26-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Thomas G. Wiseman
;       REGISTRATION NUMBER:  35046
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)861-3000
;       TELEFAX:  (202)822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  224 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-693-573-4


  Query Match             47.6%;  Score 69;  DB 18;  Length 224;
  Best Local Similarity   88.2%;  Pred. No. 4.2e-13;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147

Qy     106 GGGGGCCTAACTAACTAATTTTGTT 130
           |||||||       |   |||||||
Db     148 GGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT   7
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE:  210121.455C1
;  CURRENT APPLICATION NUMBER: US/09/123,912A
;  CURRENT FILING DATE:  1998-07-27
;  PRIOR APPLICATION NUMBER:  09/040,802
;  PRIOR FILING DATE:  1998-03-18
;  NUMBER OF SEQ ID NOS:  114
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (570)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (591)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (655)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (664)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (667)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (683)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (711)
;   OTHER INFORMATION: Where n is a, c, g or t
```

```
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (792)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-91


  Query Match             47.3%;  Score 68.6;  DB 40;  Length 858;
  Best Local Similarity   84.6%;  Pred. No. 8.2e-13;
  Matches   77;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      22 ATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGC 81
           | | | ||    ||| | |     ||||||||||||||||||||||||||||||||||
Db     333 aacgaattccaccacactggactagtggatccccgggctgcaggaattcgatatcaagc 392

Qy      82 TTATCGATACCGTCGACCTCGAGGGGGGGCC 112
           ||||||||||||||||||||||||||||||
Db     393 ttatcgataccgtcgacctcgagggggggcc 423


RESULT   8
US-09-221-107-91
; Sequence 91, Application US/09221107
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE: 210121.455C2
;  CURRENT APPLICATION NUMBER: US/09/221,107
;  CURRENT FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 161
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (570)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (591)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (655)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (664)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (667)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (683)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (711)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (792)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-91
```

```
  Query Match             47.3%;  Score 68.6;  DB 43;  Length 858;
  Best Local Similarity   84.6%;  Pred. No. 8.2e-13;
  Matches   77;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      22 ATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGC 81
           | | | ||    ||| | |     ||||||||||||||||||||||||||||||||||
Db     333 aacgaattccaccacactggactagtggatcccccgggctgcaggaattcgatatcaagc 392

Qy      82 TTATCGATACCGTCGACCTCGAGGGGGGGCC 112
           ||||||||||||||||||||||||||||||
Db     393 ttatcgataccgtcgacctcgagggggggcc 423


RESULT   9
US-09-285-479-91
; Sequence 91, Application US/09285479
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: LUNG CANCER
;  FILE REFERENCE: 210121.455C3
;  CURRENT APPLICATION NUMBER: US/09/285,479
;  CURRENT FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 172
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(858)
;   OTHER INFORMATION: n = A,T,C or G
US-09-285-479-91


  Query Match             47.3%;  Score 68.6;  DB 46;  Length 858;
  Best Local Similarity   84.6%;  Pred. No. 8.2e-13;
  Matches   77;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      22 ATCTATTTATTGCACGCGGCCGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGC 81
           | | | ||    ||| | |     ||||||||||||||||||||||||||||||||||
Db     333 aacgaattccaccacactggactagtggatcccccgggctgcaggaattcgatatcaagc 392

Qy      82 TTATCGATACCGTCGACCTCGAGGGGGGGCC 112
           ||||||||||||||||||||||||||||||
Db     393 ttatcgataccgtcgacctcgagggggggcc 423


RESULT  10
US-08-431-048A-121/c
; Sequence 121, Application US/08431048A
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  132
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Administrator, TESTA, HURWITZ &
;      ADDRESSEE:  THIBEAULT, LLP
;      STREET:  125 High Street
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,048A
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38,349
;      REFERENCE/DOCKET NUMBER:  CAN-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-431-048A-121


  Query Match             47.0%;  Score 68.2;  DB 15;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  11
US-08-496-841-121/c
; Sequence 121, Application US/08496841
;  GENERAL INFORMATION:
;    APPLICANT:  St. George-Hyslop, Peter
;    APPLICANT:  Rommens, Johanna M.
;    APPLICANT:  Fraser, Paul E.
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/496,841
;      FILING DATE:  28-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-496-841-121
```

```
  Query Match             47.0%;  Score 68.2;  DB 16;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  12
US-08-509-359-121/c
; Sequence 121, Application US/08509359
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359
;      FILING DATE:  31-JUL-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359-121


  Query Match             47.0%;  Score 68.2;  DB 17;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  13
US-08-509-359B-121/c
; Sequence 121, Application US/08509359B
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
;      STREET:  600 South Avenue West
;      CITY:  Westfield
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07090
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word, Version 6.0c
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359B
;      FILING DATE:  31-JUL-95
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Palisi, Thomas M.
;      REGISTRATION NUMBER:  36629
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 654-5000
;      TELEFAX:  (908) 654-7866
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359B-121


  Query Match             47.0%;  Score 68.2;  DB 17;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  14
US-09-124-523-121/c
; Sequence 121, Application US/09124523
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-523-121


  Query Match             47.0%;  Score 68.2;  DB 40;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  15
US-09-124-698-121/c
; Sequence 121, Application US/09124698
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,698
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-698-121


  Query Match             47.0%;  Score 68.2;  DB 40;  Length 793;
  Best Local Similarity   85.4%;  Pred. No. 1.1e-12;
  Matches   76;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      42 CGCCGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
           ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

Qy     102 GAGGGGGGGCCTAACTAACTAATTTTGTT 130
           ||||||||||||       |   |||||||
Db      62 GAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


Search completed:  May 30, 2000, 09:48:30
Job time:  60180 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:13 ; Search time 2192.43 Seconds
(without alignments)
268.067 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	70	48.3	682	96	AQ074693	AQ074693 CIT-HSP-2
c 2	68	46.9	608	96	AQ009167	AQ009167 CIT-HFP-2
c 3	67.4	46.5	528	84	B69688	B69688 CIT978SK-A-
4	67	46.2	108	84	B54292	B54292 CIT-HSP-201
5	67	46.2	141	44	AU037147	AU037147 AU037147
c 6	67	46.2	141	96	B95279	B95279 CIT-HSP-217
7	67	46.2	144	84	B80294	B80294 CIT-HSP-204
c 8	67	46.2	169	96	AQ013255	AQ013255 CIT-HSP-2
9	67	46.2	175	96	AQ041198	AQ041198 CIT-HSP-2
10	67	46.2	259	20	T48593	T48593 ph6f4_19/1T
11	67	46.2	275	27	AA052885	AA052885 T3551 Blo
c 12	67	46.2	386	84	B74162	B74162 CIT-HSP-202
c 13	67	46.2	462	96	AQ014512	AQ014512 CIT-HSP-2
c 14	67	46.2	519	84	B65534	B65534 CIT-HSP-202
15	67	46.2	547	96	AQ012629	AQ012629 CIT-HSP-2
16	67	46.2	552	96	AQ012216	AQ012216 CIT-HSP-2
17	67	46.2	575	96	AQ012221	AQ012221 CIT-HSP-2
c 18	67	46.2	596	96	AQ009165	AQ009165 CIT-HFP-2
c 19	67	46.2	607	84	B64363	B64363 CIT-HSP-202
20	67	46.2	632	84	B70299	B70299 CIT-HSP-202
c 21	67	46.2	638	96	AQ016737	AQ016737 CIT-HSP-2
22	67	46.2	647	84	B68775	B68775 CIT-HSP-205
23	67	46.2	666	84	B80296	B80296 CIT-HSP-204
24	67	46.2	700	96	AQ009153	AQ009153 CIT-HSP-2
c 25	66	45.5	541	96	AQ074695	AQ074695 CIT-HSP-2
c 26	66	45.5	576	38	AA752172	AA752172 96BS0243
27	65.4	45.1	513	96	AQ008437	AQ008437 CIT-HSP-2
c 28	65.4	45.1	541	96	AQ014519	AQ014519 CIT-HSP-2
c 29	62.8	43.3	134	84	B72088	B72088 CIT978SK-13
c 30	62.2	42.9	233	96	AQ012365	AQ012365 CIT-HSP-2
c 31	62	42.8	294	20	T26395	T26395 AB161A4F In
c 32	60.8	41.9	133	84	B26269	B26269 F10B5TF IGF
33	60	41.4	100	43	AI239372	AI239372 SWOvAFCAP
c 34	57.8	39.9	145	84	B53695	B53695 CIT-HSP-201
c 35	57.8	39.9	630	96	AQ003238	AQ003238 RPCI11-19
c 36	57.4	39.6	603	38	AA752115	AA752115 96BS0066
37	57.2	39.4	298	82	RPU68746	U68746 Riftia pach
38	56.8	39.2	97	41	AI058116	AI058116 SWAMCAC15
c 39	56	38.6	144	84	B65391	B65391 CIT-HSP-202
40	56	38.6	471	96	AQ024018	AQ024018 CpGR0359A
41	56	38.6	499	96	AQ023564	AQ023564 CpGR0482A
42	56	38.6	547	96	AQ023548	AQ023548 CpGR0051A
43	56	38.6	626	96	AQ024077	AQ024077 CpGR0415A
c 44	55.6	38.3	502	96	AQ023674	AQ023674 CpGR0007B
45	55	37.9	484	96	AQ023697	AQ023697 CpGR0111A

ALIGNMENTS

```
RESULT  1
AQ074693/c
LOCUS       AQ074693       682 bp    DNA             GSS       20-AUG-1998
DEFINITION  CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23,
            genomic survey sequence.
ACCESSION   AQ074693
VERSION     AQ074693.1  GI:3436811
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2301L23"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      156 a    187 c    187 g    151 t      1 others
ORIGIN


  Query Match           48.3%;  Score 70;  DB 96;  Length 682;
  Best Local Similarity   93.6%;  Pred. No. 1.7e-13;
  Matches   73;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 19

Qy     106 GGGGGCCTAACTAACTAA 123
           |||||||  | || | ||
Db      18 GGGGGCCCGAGTACCCAA 1


RESULT  2
AQ009167/c
LOCUS       AQ009167       608 bp    DNA             GSS       27-JUN-1998
DEFINITION  CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13,
            genomic survey sequence.
ACCESSION   AQ009167
VERSION     AQ009167.1  GI:3128599
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
            CIT-HSP-2281N13.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .608
```

```
                     /organism="Homo sapiens"
                     /db_xref="GDB:7147360"
                     /db_xref="taxon:9606"
                     /clone="2281N13"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       141 a    164 c    166 g    137 t
ORIGIN


  Query Match             46.9%;  Score 68;  DB 96;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     106 GGGGGCCT 113
           ||||||||
Db      22 GGGGGCCT 15


RESULT   3
B69688/c
LOCUS       B69688         528 bp    DNA             GSS       18-JUN-1998
DEFINITION  CIT978SK-A-448E10.TVC CIT978SK Homo sapiens genomic clone A-448E10,
            genomic survey sequence.
ACCESSION   B69688
VERSION     B69688.1  GI:2708912
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /db_xref="GDB:5299607"
                     /db_xref="taxon:9606"
                     /clone="A-448E10"
                     /clone_lib="CIT978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
BASE COUNT       118 a    138 c    147 g    125 t
ORIGIN


  Query Match             46.5%;  Score 67.4;  DB 84;  Length 528;
  Best Local Similarity   98.6%;  Pred. No. 1.3e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 16

Qy     106 GGGGGCCTA 114
           ||||||| |
Db      15 GGGGGCCCA 7


RESULT   4
B54292
LOCUS       B54292         108 bp    DNA             GSS       20-JUN-1998
DEFINITION  CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
            genomic survey sequence.
ACCESSION   B54292
VERSION     B54292.1  GI:2608626
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7044003"
                     /db_xref="taxon:9606"
                     /clone="2017M1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT        20 a     35 c     34 g     19 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 84;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 93

Qy     106 GGGGGCC 112
           |||||||
Db      94 GGGGGCC 100


RESULT   5
AU037147
```

```
LOCUS       AU037147      141 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSB532, mRNA sequence.
ACCESSION   AU037147
VERSION     AU037147.1  GI:3983900
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2153012.
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSB532"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       26 a     40 c     42 g     28 t      5 others
ORIGIN
```

```
  Query Match             46.2%;  Score 67;  DB 44;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 128

Qy     106 GGGGGCC 112
           |||||||
Db     129 GGGGGCC 135
```

```
RESULT   6
B95279/c
LOCUS       B95279        141 bp    DNA             GSS       26-JUN-1998
DEFINITION  CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
            genomic survey sequence.
ACCESSION   B95279
VERSION     B95279.1  GI:2977616
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Homo sapiens"
                     /db_xref="GDB:7103988"
                     /db_xref="taxon:9606"
                     /clone="2172N2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       24 a     41 c     41 g     35 t
ORIGIN
```

```
  Query Match             46.2%;  Score 67;  DB 96;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 14

Qy     106 GGGGGCC 112
           |||||||
Db      13 GGGGGCC 7
```

```
RESULT   7
B80294
LOCUS       B80294        144 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
            genomic survey sequence.
ACCESSION   B80294
VERSION     B80294.1  GI:2867317
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2045D19.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="Homo sapiens"
                     /db_xref="GDB:7054558"
                     /db_xref="taxon:9606"
```

```
                     /clone="2045D19"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       34 a     43 c     41 g     26 t
ORIGIN


  Query Match           46.2%;  Score 67;  DB 84;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 129

Qy     106 GGGGGCC 112
           |||||||
Db     130 GGGGGCC 136


RESULT   8
AQ013255/c
LOCUS       AQ013255      169 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
            genomic survey sequence.
ACCESSION   AQ013255
VERSION     AQ013255.1  GI:3185820
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2299C22.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .169
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154018"
                     /db_xref="taxon:9606"
                     /clone="2299C22"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       32 a     47 c     48 g     42 t
ORIGIN


  Query Match           46.2%;  Score 67;  DB 96;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 42

Qy     106 GGGGGCC 112
           |||||||
Db      41 GGGGGCC 35


RESULT   9
AQ041198
LOCUS       AQ041198      175 bp    DNA             GSS       14-JUL-1998
DEFINITION  CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
            genomic survey sequence.
ACCESSION   AQ041198
VERSION     AQ041198.1  GI:3310469
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .175
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2335D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       43 a     50 c     47 g     35 t
ORIGIN


  Query Match           46.2%;  Score 67;  DB 96;  Length 175;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132

Qy     106 GGGGGCC 112
           |||||||
Db     133 GGGGGCC 139


RESULT  10
T48593
LOCUS       T48593        259 bp    mRNA            EST       02-FEB-1995
DEFINITION  ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph6f4_19/1TV, mRNA sequence.
```

```
ACCESSION   T48593
VERSION     T48593.1  GI:642793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L.,
            Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
            Sverdlov,E.D.
  TITLE     Outward Alu-primed hncDNA library
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Sverdlov ED
            Structure and Function of Human Genes
            Shemyakin Institute of Bioorganic Chemistry
            16/10 Miklukho-Maklaya, Moscow, 117871, Russia
            Tel: 70953306529
            Fax: 70953306538
            Email: sverd@humgen.siobc.msk.su.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ph6f4_19/1TV"
                     /clone_lib="Outward Alu-primed hncDNA library"
                     /note="Vector: pGEM-3Z; Site_1: EcoR1; Site_2: BamH1; The
                     library was constructed as described in [Obradovic, D.,
                     Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
                     Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
                     Sverdlov, E.D. (1993) Bioorganicheskaya khimia, 20,
                     919-930]. This protocol is based on nested primer strategy
                     using Alu- specific primers (ALN3 and TC-65) that direct
                     the hncDNA synthesis outward of Alu repeats."
BASE COUNT       62 a     66 c     72 g     59 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 20;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 115

Qy     106 GGGGGCC 112
           |||||||
Db     116 GGGGGCC 122


RESULT  11
AA052885
LOCUS       AA052885      275 bp    mRNA            EST       13-SEP-1996
DEFINITION  T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
            brucei cDNA 5', mRNA sequence.
ACCESSION   AA052885
VERSION     AA052885.1  GI:1543898
KEYWORDS    EST.
SOURCE      Trypanosoma brucei brucei.
  ORGANISM  Trypanosoma brucei brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Osanya,A., Murphy,N.B. and Pelle,R.
  TITLE     Trypanosoma brucei brucei cDNAs
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:637578.
            Contact: Osanya A
            Unit 3
            International Livestock Research Institute
            Box 30709, Nairobi, Kenya
            Tel: 254 2 630 743
            Fax: 254 2 631 499
            Email: a.osanya@cgnet.com
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Trypanosoma brucei brucei"
                     /strain="Clone A4"
                     /db_xref="taxon:5702"
                     /clone_lib="Bloodstream form of serodeme ILTat1.1"
                     /note="cDNAs were generated from poly (A+) enriched mRNA
                     prepared from different developmental stages of T.b.brucei
                     by reverse transcription followed by PCR amplification
                     using mini-exon and oligo(dT) primers.  The cDNA generated
                     were utilized in RADES-PCR coupled with differential
                     hybridisations to identify differentially expressed mRNA
                     transcripts. The products which showed to be
                     differentially expressed were cloned pGEM -T vector.
                     These differentially expressed mRNA transcripts were (are
                     being) sequenced to generate differentially expressed
                     sequence tags."
BASE COUNT       61 a     84 c     74 g     56 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 27;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 108

Qy     106 GGGGGCC 112
           |||||||
Db     109 GGGGGCC 115


RESULT  12
B74162/c
LOCUS       B74162        386 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
            genomic survey sequence.
ACCESSION   B74162
VERSION     B74162.1  GI:2769849
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 386)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2028N1.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .386
                     /organism="Homo sapiens"
                     /db_xref="GDB:7048251"
```

```
                     /db_xref="taxon:9606"
                     /clone="2028N1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       81 a     98 c    109 g     98 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 84;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     106 GGGGGCC 112
           |||||||
Db      22 GGGGGCC 16


RESULT  13
AQ014512/c
LOCUS       AQ014512      462 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
            genomic survey sequence.
ACCESSION   AQ014512
VERSION     AQ014512.1  GI:3184839
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154425"
                     /db_xref="taxon:9606"
                     /clone="2300D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       89 a    127 c    131 g    115 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 96;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 40

Qy     106 GGGGGCC 112
           |||||||
Db      39 GGGGGCC 33


RESULT  14
B65534/c
LOCUS       B65534        519 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
            genomic survey sequence.
ACCESSION   B65534
VERSION     B65534.1  GI:2639512
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /db_xref="GDB:7045730"
                     /db_xref="taxon:9606"
                     /clone="2022D24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      110 a    137 c    148 g    124 t
ORIGIN


  Query Match             46.2%;  Score 67;  DB 84;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     106 GGGGGCC 112
           |||||||
Db      22 GGGGGCC 16


RESULT  15
AQ012629
LOCUS       AQ012629      547 bp    DNA             GSS       06-JUN-1998
```

```
DEFINITION  CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
            genomic survey sequence.
ACCESSION   AQ012629
VERSION     AQ012629.1  GI:3185194
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154069"
                     /db_xref="taxon:9606"
                     /clone="2299F1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      154 a    139 c    125 g    129 t
ORIGIN
```

Query Match 46.2%; Score 67; DB 96; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.8e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      46 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 134

Qy     106 GGGGGCC 112
           |||||||
Db     135 GGGGGCC 141
```

Search completed: May 29, 2000, 21:13:15
Job time: 36514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:29 ; Search time 1214.87 Seconds
(without alignments)
927.949 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	69.8	47.2	259	5	A44281	A44281 Sequence 9
	2	69.4	46.9	2958	14	ARBLKSM	X52326 pBluescript
	3	69.4	46.9	2958	14	ARBLKSP	X52331 pBluescript
	4	69.4	46.9	2961	14	ARBL2KSM	X52329 pBluescript
	5	69.4	46.9	2961	14	ARBL2KSP	X52327 pBluescript
	6	69.4	46.9	2964	14	SYNBLKSMV	L08784 BlueScribe
	7	69.4	46.9	2964	14	SYNBLKSPV	L08785 BlueScribe
c	8	69.4	46.9	3306	14	SYNPBEN66	D85525 Cloning vec
	9	69.4	46.9	4144	14	XXU35131	U35131 Plasmid pBS
	10	69.4	46.9	4267	14	PRS304	U03436 Yeast integ
	11	69.4	46.9	4289	14	XXU35136	U35136 Plasmid pBS
	12	69.4	46.9	4373	14	PRS306	U03438 Yeast integ
	13	69.4	46.9	4443	14	PRS303	U03435 Yeast integ
c	14	69.4	46.9	4670	14	ASAJ5326	AJ005326 pGAII(+)
	15	69.4	46.9	4670	14	ASAJ5329	AJ005329 pGAII(-)
	16	69.4	46.9	4707	14	XXU02374	U02374 Cloning vec
	17	69.4	46.9	4768	14	XXU25061	U25061 Cloning vec
	18	69.4	46.9	4783	14	PRS314	U03440 Yeast centr
	19	69.4	46.9	4887	14	PRS316	U03442 Yeast centr
	20	69.4	46.9	4950	14	XXU25060	U25060 Cloning vec
	21	69.4	46.9	4967	14	PRS313	U03439 Yeast centr
	22	69.4	46.9	5144	14	CVU23751	U23751 Cloning vec
	23	69.4	46.9	5187	14	U34887	U34887 Yeast integ
	24	69.4	46.9	5228	14	XXU25059	U25059 Cloning vec
	25	69.4	46.9	5504	14	PRS305	U03437 Yeast integ
	26	69.4	46.9	5634	14	CVU14125	U14125 Cloning vec
	27	69.4	46.9	6018	14	PRS315	U03441 Yeast centr
c	28	69.4	46.9	6340	14	ASAJ5323	AJ005323 pCP1(-) K
	29	69.4	46.9	9655	14	SYNPGR8V	AD001531 Cloning v
c	30	68.8	46.5	2961	14	CVU46017	U46017 Cloning vec
c	31	68.8	46.5	3240	8	AF015771	AF015771 Magnaport
c	32	68.8	46.5	3399	14	CVU46018	U46018 Cloning vec
	33	68.8	46.5	4278	14	AF028239	AF028239 Mammalian
	34	68.8	46.5	4307	14	AF067646	AF067646 Cloning v
	35	68.6	46.4	4328	14	AF072999	AF072999 Cloning v
c	36	68.4	46.2	793	5	AR060142	AR060142 Sequence
	37	68.4	46.2	4323	14	AF072538	AF072538 Cloning v
	38	68.4	46.2	4324	14	AF072539	AF072539 Cloning v
	39	68.4	46.2	4325	14	AF072540	AF072540 Cloning v
	40	68.4	46.2	4326	14	AF072997	AF072997 Cloning v
	41	68.4	46.2	4327	14	AF072998	AF072998 Cloning v
	42	68.2	46.1	3727	14	SYNDROCL	M32616 Synthetic c
c	43	68	45.9	338	1	BSPX91477	X91477 Bacterial s
c	44	68	45.9	6037	14	YSCTRAM2	M74016 Cloning vec
c	45	68	45.9	6445	14	SYNECOYST	L11060 Cloning vec

ALIGNMENTS

```
RESULT   1
A44281/c
LOCUS       A44281         259 bp    DNA              PAT       07-MAR-1997
DEFINITION  Sequence 9 from Patent WO9511300.
ACCESSION   A44281
VERSION     A44281.1  GI:2299115
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Chandley,A.C., Kun,M., Sharkey,A.M., Hargreave,T.B. and Cooke,H.J.
  TITLE     AZOOSPERMIA IDENTIFICATION AND TREATMENT
  JOURNAL   Patent: WO 9511300-A 9 27-APR-1995;
            MEDICAL RES COUNCIL (GB)
COMMENT     Other publication AU 7947794 950508.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /clone_lib="MOUSE GENOMIC LIBRARY"
                     /clone="M3.2"
BASE COUNT       67 a     58 c     63 g     69 t      2 others
ORIGIN

  Query Match             47.2%;  Score 69.8;  DB 5;  Length 259;
  Best Local Similarity   79.0%;  Pred. No. 9.4e-12;
  Matches   83;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      29 TATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT 88
           ||  | | ||      |  | ||||||||||||||||||||||||||||||||||||||||
Db     132 TAAGGAAAGCAAAGCCTCTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT 73

Qy      89 CGATACCGTCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||||||||||||||||||||||      |   |||||||
Db      72 CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 28


RESULT   2
ARBLKSM
LOCUS       ARBLKSM      2958 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52326
VERSION     X52326.1  GI:58064
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(-)"
BASE COUNT      708 a    754 c    731 g    765 t
ORIGIN

  Query Match             46.9%;  Score 69.4;  DB 14;  Length 2958;
  Best Local Similarity   87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   3
ARBLKSP
LOCUS       ARBLKSP      2958 bp    DNA   circular  SYN       11-JUN-1998
DEFINITION  pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52331
VERSION     X52331.1  GI:58065
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
REFERENCE   4  (bases 1 to 2958)
  AUTHORS   Lampe,D.J., Grant,T.E. and Robertson,H.M.
  TITLE     Factors affecting transposition of the Himar1 mariner transposon in
            vitro
  JOURNAL   Genetics 149 (1), 179-187 (1998)
  MEDLINE   98250682
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(+)"
BASE COUNT      749 a    734 c    751 g    724 t
ORIGIN

  Query Match             46.9%;  Score 69.4;  DB 14;  Length 2958;
  Best Local Similarity   87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   4
```

```
ARBL2KSM
LOCUS       ARBL2KSM     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(-) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52329
VERSION     X52329.1  GI:58060
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(-)"
BASE COUNT      706 a    758 c    735 g    762 t
ORIGIN


  Query Match            46.9%;  Score 69.4;  DB 14;  Length 2961;
  Best Local Similarity  87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   5
ARBL2KSP
LOCUS       ARBL2KSP     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(+) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52327
VERSION     X52327.1  GI:58061
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(+)"
BASE COUNT      747 a    738 c    755 g    721 t
ORIGIN


  Query Match            46.9%;  Score 69.4;  DB 14;  Length 2961;
  Best Local Similarity  87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   6
SYNBLKSMV
LOCUS       SYNBLKSMV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Minus cloning vector.
ACCESSION   L08784
VERSION     L08784.1  GI:310728
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Minus - Cloning vector
            ENTRY BLUEKSM                  #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Minus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0077
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the minus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
```

```
            #parent
              VecBase(3):BlueM13m
            #parent
              VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
              VecSource(3):bGalKS, GenBank(50):PF1
            #brother
              VecBase(3):BlueKSp, VecBase(3):BlueSKm
         PARENT
            Features of BlueKSm (2964 bp)
              residue     source
               3- 458  5488-5943     phage f1
              460- 624   236- 400     pUC19c
              626- 645     1-  20     T7 promoter
              653- 760   108-   1 (c) BlueKS-polylinker
              772- 791    20-   1 (c) T3 promoter
              795-2964   448-2617     pUC19c
            Conflict (cfl) and Mutations (mut): none
         PARENT
            Features of BlueKSm (2964 bp)
              residue     source
               3- 458  5488-5943     phage f1
              460- 624   449- 285 (c) pUC19
              626- 645     1-  20     T7 promoter
              653- 760   108-   1 (c) BlueKS-polylinker
              772- 791    20-   1 (c) T3 promoter
              795-1031   237-   1 (c) pUC19
             1032-2964  2686- 754 (c) pUC19
            Conflict (cfl) and Mutations (mut): none
         FEATURE
                    643          start of T7-RNA synthesis
                    774       (c) start of T3-RNA synthesis
            1976-2764 789-1 (c) Ap-R; b-lactamase
         POLYLINKER
              KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
              SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
            #resistance  Ap
            #indicator beta-galactosidase
         SUMMARY  BlueKSm    #length 2964   #checksum 1589.
FEATURES             Location/Qualifiers
     source          1. .2964
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      708 a    756 c    735 g    765 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 2964;
  Best Local Similarity   87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   7
SYNBLKSPV
LOCUS       SYNBLKSPV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Plus cloning vector.
ACCESSION   L08785
VERSION     L08785.1  GI:310729
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Plus - Cloning vector
            ENTRY BLUEKSP                 #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Plus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0078
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the plus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
               #parent
                 VecBase(3):BlueM13p
               #parent
                 VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                 VecSource(3):bGalKS, GenBank(50):PF1
               #brother
                 VecBase(3):BlueKSm, VecBase(3):BlueSKp
            PARENT
               Features of BlueKSp (2964 bp)
                 residue     source
                  3- 458  5943-5488 (c) phage f1
                 460- 624   236- 400     pUC19c
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-2964   448-2617     pUC19c
               Conflict (cfl) and Mutations (mut): none
            PARENT
               Features of BlueKSp (2964 bp)
                 residue     source
                  3- 458  5943-5488 (c) phage f1
                 460- 624   449- 285 (c) pUC19
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-1031   237-   1 (c) pUC19
                1032-2964  2686- 754 (c) pUC19
               Conflict (cfl) and Mutations (mut): none
            FEATURE
                     643          start of T7-RNA synthesis
                     774       (c) start of T3-RNA synthesis
               1976-2764 789-1 (c) Ap-R; b-lactamase
            POLYLINKER
                 KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
                 SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
               #resistance  Ap
               #indicator beta-galactosidase
            SUMMARY  BlueKSp    #length 2964   #checksum 690.
FEATURES             Location/Qualifiers
     source          1. .2964
```

```
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      750 a    736 c    755 g    723 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 2964;
  Best Local Similarity   87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     745 GGGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   8
SYNPBEN66/c
LOCUS       SYNPBEN66     3306 bp    DNA   circular  SYN       06-FEB-1999
DEFINITION  Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
            beta-lactamase, complete cds.
ACCESSION   D85525
VERSION     D85525.1  GI:1345433
KEYWORDS    plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE      Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
  ORGANISM  Cloning vector pBEN66
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 3306)
  AUTHORS   Yamamoto,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
            Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
            Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
            (Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE   2  (sites)
  AUTHORS   Yamamoto,Y. and Furuyama,J.
  TITLE     One-step disruption by circular DNA in Escherichia coli
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1. .3306
                     /organism="Cloning vector pBEN66"
                     /plasmid="pBEN66"
                     /db_xref="taxon:47800"
                     /lab_host="E.coli"
     promoter        19. .36
                     /note="T3 promoter"
     gene            260. .1075
                     /gene="kan"
     CDS             260. .1075
                     /gene="kan"
                     /note="derived from Tn903; kanamycin resistance gene"
                     /citation=[2]
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside 3'-phosphotransferase"
                     /protein_id="BAA12824.1"
                     /db_xref="GI:1345434"
                     /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                     GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                     GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                     VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
                     RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
     gene            complement(2209. .3069)
                     /gene="amp"
     CDS             complement(2209. .3069)
                     /gene="amp"
                     /note="amipicillin resistance gene"
                     /citation=[1]
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="BAA12825.1"
                     /db_xref="GI:1345435"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT      854 a    800 c    790 g    862 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 3306;
  Best Local Similarity   87.4%;  Pred. No. 1.6e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 64

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db      63 GGGGGGGCCCGGTACCCAGCTTTTGTT 37


RESULT   9
XXU35131
LOCUS       XXU35131     4144 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL159 cloning vector, complete sequence.
ACCESSION   U35131
VERSION     U35131.1  GI:984907
KEYWORDS    .
SOURCE      cloning vector pBSL159.
  ORGANISM  Cloning vector pBSL159
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4144)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4144)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4144
                     /organism="Cloning vector pBSL159"
                     /plasmid="pBSL159"
                     /db_xref="taxon:42704"
     CDS             complement(1129. .1989)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53618.1"
                     /db_xref="GI:984908"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     CDS             complement(join(3704. .4144,1. .219))
                     /codon_start=1
                     /transl_table=11
                     /product="chloramphenicol acetyltransferase"
                     /protein_id="AAC53619.1"
                     /db_xref="GI:984909"
```

```
                     /translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAF
                     LKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETF
                     SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNV
                     ANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGG
                     A"
BASE COUNT     1084 a   1009 c    959 g   1092 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 4144;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2724 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2783

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db    2784 GGGGGGGCCCGGTACCCAGCTTTTGTT 2810


RESULT  10
PRS304
LOCUS       PRS304       4267 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS304 with TRP1 marker, complete
            sequence.
ACCESSION   U03436
VERSION     U03436.1  GI:416305
KEYWORDS    .
SOURCE      Cloning vector pRS304.
  ORGANISM  Cloning vector pRS304
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4267)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4267)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4267
                     /organism="Cloning vector pRS304"
                     /db_xref="taxon:31827"
BASE COUNT     1128 a    965 c   1092 g   1082 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 4267;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1919 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 1978

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db    1979 GGGGGGGCCCGGTACCCAGCTTTTGTT 2005


RESULT  11
XXU35136
LOCUS       XXU35136     4289 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION   U35136
VERSION     U35136.1  GI:984923
KEYWORDS    .
SOURCE      cloning vector pBSL97.
  ORGANISM  Cloning vector pBSL97
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4289)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4289)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4289
                     /organism="Cloning vector pBSL97"
                     /plasmid="pBSL97"
                     /db_xref="taxon:42709"
     CDS             complement(804. .1598)
                     /codon_start=1
                     /transl_table=11
                     /product="neomycin phosphotransferase"
                     /protein_id="AAC53629.1"
                     /db_xref="GI:984924"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
     CDS             complement(3301. .4161)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53630.1"
                     /db_xref="GI:984925"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     1024 a   1157 c   1111 g    997 t
ORIGIN


  Query Match             46.9%;  Score 69.4;  DB 14;  Length 4289;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2013 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2072

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db    2073 GGGGGGGCCCGGTACCCAGCTTTTGTT 2099


RESULT  12
PRS306
LOCUS       PRS306       4373 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS306 with URA3 marker, complete
            sequence.
ACCESSION   U03438
VERSION     U03438.1  GI:416307
KEYWORDS    .
SOURCE      Cloning vector pRS306.
  ORGANISM  Cloning vector pRS306
            artificial sequence; vectors.
```

```
REFERENCE   1  (bases 1 to 4373)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4373)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4373
                     /organism="Cloning vector pRS306"
                     /db_xref="taxon:31829"
BASE COUNT     1181 a    987 c   1106 g   1099 t
ORIGIN


  Query Match           46.9%;  Score 69.4;  DB 14;  Length 4373;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2025 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2084

Qy    107 GGGGGGGCCTAACTAACTAATTTTGTT 133
          |||||||||       |   |||||||
Db   2085 GGGGGGGCCCGGTACCCAGCTTTTGTT 2111


RESULT  13
PRS303
LOCUS       PRS303       4443 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS303 with HIS3 marker, complete
            sequence.
ACCESSION   U03435
VERSION     U03435.1  GI:416304
KEYWORDS    .
SOURCE      Cloning vector pRS303.
  ORGANISM  Cloning vector pRS303
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4443)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4443)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4443
                     /organism="Cloning vector pRS303"
                     /db_xref="taxon:31826"
BASE COUNT     1149 a   1048 c   1109 g   1137 t
ORIGIN


  Query Match           46.9%;  Score 69.4;  DB 14;  Length 4443;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2105 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2164

Qy    107 GGGGGGGCCTAACTAACTAATTTTGTT 133
          |||||||||       |   |||||||
Db   2165 GGGGGGGCCCGGTACCCAGCTTTTGTT 2191


RESULT  14
ASAJ5326/c
LOCUS       ASAJ5326     4670 bp    DNA   circular  SYN       08-FEB-1999
DEFINITION  pGAII(+) KS positive selection cloning vector gltS gene.
ACCESSION   AJ005326
VERSION     AJ005326.1  GI:4028984
KEYWORDS    gltS gene; glutamate permease.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4670)
  AUTHORS   Gal,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
            Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
            6726, HUNGARY
REFERENCE   2  (bases 1 to 4670)
  AUTHORS   Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
            Kalman,M.
  TITLE     A positive selection cloning system based on the gltS gene of
            Escherichia coli
  JOURNAL   Anal. Biochem. 266 (2), 235-238 (1999)
  MEDLINE   99107575
FEATURES             Location/Qualifiers
     source          1. .4670
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     gene            852. .2231
                     /gene="gltS"
     CDS             852. .2231
                     /gene="gltS"
                     /codon_start=1
                     /transl_table=11
                     /product="glutamate permease"
                     /protein_id="CAA06473.1"
                     /db_xref="GI:4028985"
                     /translation="MFHLDTLATLVAATLTLLLGRKLVHSVSFLKKYTIPEPVAGGLL
                     VALALLVLKKSMGWEVNFDMSLRDPLMLAFFATIGLNANIASLRAGGRVVGIFLIVVV
                     GLLVMQNAIGIGMASLLGLDPLMGLLAGSITLSGGHGTGAAWSKLFIERYGFTNATEV
                     AMACATFGLVLGGLIGGPVARQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGDPLV
                     LERPPPRWSSNSPYSESYYARYLVKHSTTPNGIPDDQEVPTAFEKPDVGRMITSLVLI
                     ETIALIAICLTVGKIVAQLLAGTAFELPTFVCVLFVGVILSNGLSIMGFYRVFERAVS
                     VLGNVSLSLFLAMALMGLKLWELASLALPMLAILVVQTIFMALYAIFVTWRMMGKNYD
                     AAVLAAGHCGFGLGATPTAIANMQAITERFGPSHMAFLVVPMVGAFFIDIVNALVIKL
                     YLMLPIFAG"
BASE COUNT     1040 a   1165 c   1231 g   1234 t
ORIGIN


  Query Match           46.9%;  Score 69.4;  DB 14;  Length 4670;
  Best Local Similarity   87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1502 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 1443

Qy    107 GGGGGGGCCTAACTAACTAATTTTGTT 133
          |||||||||       |   |||||||
Db   1442 GGGGGGGCCCGGTACCCAGCTTTTGTT 1416


RESULT  15
ASAJ5329
LOCUS       ASAJ5329     4670 bp    DNA   circular  SYN       08-FEB-1999
DEFINITION  pGAII(-) KS positive selection cloning vector gltS gene.
ACCESSION   AJ005329
VERSION     AJ005329.1  GI:4028990
KEYWORDS    gltS gene; glutamate permease.
```

```
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4670)
  AUTHORS   Gal,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
            Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
            6726, HUNGARY
REFERENCE   2  (bases 1 to 4670)
  AUTHORS   Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
            Kalman,M.
  TITLE     A positive selection cloning system based on the gltS gene of
            Escherichia coli
  JOURNAL   Anal. Biochem. 266 (2), 235-238 (1999)
  MEDLINE   99107575
FEATURES             Location/Qualifiers
     source          1. .4670
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     gene            complement(1150. .2529)
                     /gene="gltS"
     CDS             complement(1150. .2529)
                     /gene="gltS"
                     /codon_start=1
                     /transl_table=11
                     /product="glutamate permease"
                     /protein_id="CAA06476.1"
                     /db_xref="GI:4028991"
                     /translation="MFHLDTLATLVAATLTLLLGRKLVHSVSFLKKYTIPEPVAGGLL
                     VALALLVLKKSMGWEVNFDMSLRDPLMLAFFATIGLNANIASLRAGGRVVGIFLIVVV
                     GLLVMQNAIGIGMASLLGLDPLMGLLAGSITLSGGHGTGAAWSKLFIERYGFTNATEV
                     AMACATFGLVLGGLIGGPVARQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGDPLV
                     LERPPPRWSSNSPYSESYYARYLVKHSTTPNGIPDDQEVPTAFEKPDVGRMITSLVLI
                     ETIALIAICLTVGKIVAQLLAGTAFELPTFVCVLFVGVILSNGLSIMGFYRVFERAVS
                     VLGNVSLSLFLAMALMGLKLWELASLALPMLAILVVQTIFMALYAIFVTWRMMGKNYD
                     AAVLAAGHCGFGLGATPTAIANMQAITERFGPSHMAFLVVPMVGAFFIDIVNALVIKL
                     YLMLPIFAG"
BASE COUNT     1198 a   1251 c   1145 g   1076 t
ORIGIN


  Query Match           46.9%;  Score 69.4;  DB 14;  Length 4670;
  Best Local Similarity  87.4%;  Pred. No. 1.7e-11;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1879 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 1938

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db    1939 GGGGGGGCCCGGTACCCAGCTTTTGTT 1965


Search completed: May 29, 2000, 21:35:34
Job time: 36866 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:21 ; Search time 1446.77 Seconds
(without alignments)
25.594 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	71.8	48.5	10529	1	V09028	Maize fluory2 gene
c	2	69.8	47.2	259	1	Q87664	Mouse azoospermia
	3	69.8	47.2	501	1	T04866	Nucleotide analogu
c	4	68.6	46.4	1091	1	T75006	Human endogenous r
c	5	68	45.9	5356	1	T43794	Plasmid pRIPHAT (r
c	6	67.6	45.7	1122	1	T75005	Human endogenous r
	7	67.4	45.5	651	1	X20513	Polynucleotide seq
c	8	67.4	45.5	752	1	V31294	E. coli J96 pathog
c	9	67.4	45.5	3792	1	Q48463	Plasmid pG+host4 c
	10	67.4	45.5	4226	1	T39485	Human steroidogene
c	11	67.4	45.5	5234	1	Q48464	Plasmid pG+host5 c
c	12	67.4	45.5	6722	1	Q48465	Plasmid pG+host6 c
	13	67.4	45.5	12814	1	X24730	Swedish-FAD APP ta
	14	67.4	45.5	15692	1	X24731	London-FAD APP tar
	15	67.4	45.5	15692	1	X24732	Swedish/London-FAD
	16	67.4	45.5	15701	1	X24733	Swedish-FAD APP713
c	17	67.2	45.4	545	1	V68808	Human endogenous r
c	18	65.4	44.2	2640	1	Q26664	bDAT. cDNA encodin
	19	65	43.9	84	1	X02800	E. coli biotin DNA
c	20	65	43.9	3465	1	X02814	DE19731274 Seq ID
c	21	65	43.9	3481	1	X02815	DE19731274 Seq ID
c	22	65	43.9	3794	1	X02812	DE19731274 Seq ID
c	23	65	43.9	3810	1	X02813	DE19731274 Seq ID
	24	65	43.9	6596	1	V57377	Maize female-prefe
c	25	64	43.2	545	1	T75010	Human endogenous r
	26	62.8	42.4	685	1	V10190	Stealth virus nucl
	27	62.8	42.4	685	1	V12003	Stealth virus plas
	28	59.6	40.3	4164	1	T04575	Plasmid pAT-1 sequ
	29	59.6	40.3	4164	1	V22271	pAT-1 (pSD544). DN
	30	59.6	40.3	4164	1	V69740	Nucleotide sequenc
	31	58.4	39.5	5178	1	T49876	pTet-Splice. Nucle
	32	58.2	39.3	3198	1	T92702	Candida CaRhol gen
	33	58.2	39.3	3198	1	T92869	Candida CaRhol gen
	34	56.4	38.1	459	1	V57250	Clone #4 from muta
	35	56.4	38.1	1949	1	X05602	Nucleotide sequenc
c	36	56	37.8	78	1	X02799	E. coli biotin DNA
c	37	56	37.8	2150	1	T91037	Yeast checkpoint c
c	38	56	37.8	2150	1	X01271	Yeast RAD17 coding
	39	56	37.8	2973	1	V64254	Plasmid pKS varian
	40	56	37.8	3956	1	V64258	Plasmid pPK13/14 D
	41	56	37.8	4088	1	V64255	Plasmid pPK5/6 DNA
	42	56	37.8	4102	1	V64257	Plasmid pPK9/10 DN
	43	56	37.8	4583	1	V64256	Plasmid pPK7/8 DNA
	44	54.4	36.8	10504	1	V79503	DNA sequence of ex
	45	54	36.5	1094	1	X29498	S. aureus MurF ORF

ALIGNMENTS

RESULT 1
V09028/c
ID V09028 standard; DNA; 10529 BP.
AC V09028;
DT 20-JUL-1998 (first entry)
DE Maize fluory2 gene (f12).
KW Floury2; f12 gene; alpha-zein; signal peptide; transgenic plant;
KW seed; essential amino acid; animal feed; maize; rice; wheat;
KW barley; millet; sorghum; ds.
OS Zea mays.
FH Key Location/Qualifiers
FT misc_feature 1. .760
FT /*tag= a
FT /note= "vector-derived sequence"
FT promoter 761. .3824
FT /*tag= b
FT /note= "Claim 14"
FT CDS 3825. .4613
FT /*tag= c
FT sig_peptide 3825. .3887
FT /*tag= d
FT mat_peptide 3888. .4610
FT /*tag= e
PN WO9802563-A1.
PD 22-JAN-1998.
PF 11-JUL-1997; U11723.
PR 17-JUL-1996; US-021833.
PA (PION-) PIONEER HI-BRED INT INC.
PA (UYAR-) UNIV ARIZONA STATE.
PI Beach L, Coleman CE, Larkins BA;
DR WPI; 98-110609/10.
DR P-PSDB; W23977.
PT Cereal plants containing trans-gene expressing fusion that includes
PT signal peptide of the f12 maize gene - and protein having high
PT content of essential amino acids, producing feeds of improved
PT nutritional value
PS Claim 14; Fig 1A-H; 37pp; English.
CC This is the nucleotide sequence of a clone of the fluory2 (f12)
CC gene of maize. It codes for a 24-kDa alpha-zein protein (see
CC W23977) that includes a 21-amino acid signal peptide (see W23976)
CC which targets the alpha-zein to the lumen of the rough endoplasmic
CC reticulum. A claimed cereal plant contains a transgene comprising
CC a first polynucleotide that encodes the f12 signal peptide and a
CC second polynucleotide that encodes an agronomically high-value
CC protein. Also new are seeds produced by the plants, the transgene
CC itself, and a transgene that also includes the f12 promoter. The
CC second polynucleotide preferably encodes a protein that has a high
CC content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC maize, wheat, rice, barley, millet or sorghum will have increased
CC contents of these essential amino acids in their seeds (all
CC claimed).
SQ Sequence 10529 BP; 2845 A; 2484 C; 2262 G; 2938 T;

Query Match 48.5%; Score 71.8; DB 1; Length 10529;
Best Local Similarity 76.5%; Pred. No. 9.8e-16;

```
  Matches   88;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCG 60
           ||| |||| |  |  | ||  | |    |   |  |||    |   | ||||||||||||
Db     774 GGCTAAAATTGCATGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCG 715

Qy      61 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     714 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 660


RESULT   2
Q87664/c
ID   Q87664 standard; DNA; 259 BP.
AC   Q87664;
DT   06-MAR-1996  (first entry)
DE   Mouse azoospermia factor (AZF) gene partial clone M3.2.
KW   Azoospermia factor; AZF; male infertility; YRRM gene;
KW   Y-chromosome; RNA recognition motif; ds.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   misc_difference 174
FT                   /*tag=  a
FT                   /note= "base n at position 174 is not identified
FT                   in the specification"
FT   misc_difference 197
FT                   /*tag=  b
FT                   /note= "base n at position 197 is not identified
FT                   in the specification"
PN   WO9511300-A2.
PD   27-APR-1995.
PF   24-OCT-1994; G02344.
PR   22-OCT-1993; GB-021857.
PR   07-JUL-1994; GB-013760.
PA   (MEDI-) MEDICAL RES COUNCIL.
PI   Chandley AC,  Cooke HJ,  Hargreave TB,  Kun M,  Sharkey AM;
DR   WPI; 95-170221/22.
PT   Nucleic acid encoding the human azoospermia factor, and probes and
PT   antibodies specific for the sequence and encoded polypeptide - may
PT   be used in the clinical diagnosis of male infertility
PS   Disclosure; Fig 5; 40pp; English.
CC   A mouse gene partial sequence (Q87664) showing homology to the human
CC   azoospermia factor (AZF) YRMM gene was obtd. from mouse genomic phage
CC   clone M3.2 isolated from a genomic library screened with MYK (YRRM1)
CC   cDNA (Q87655).  M3.2, cloned in Lambda 2001, was mapped to the short arm
CC   of the mouse Y-chromosome.  A full cDNA sequence (Q87665) was also
CC   obtd.
SQ   Sequence   259 BP;    67 A;    58 C;    63 G;    69 T;


  Query Match             47.2%;  Score 69.8;  DB 1;  Length 259;
  Best Local Similarity   79.0%;  Pred. No. 1.4e-15;
  Matches   83;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      29 TATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT 88
           ||  | | ||     |  | ||||||||||||||||||||||||||||||||||||||||
Db     132 TAAGGAAAGCAAAGCCTCTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTAT 73

Qy      89 CGATACCGTCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||||||||||||||||||       |   |||||||
Db      72 CGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 28


RESULT   3
T04866
ID   T04866 standard; DNA; 501 BP.
AC   T04866;
DT   28-JAN-1996  (first entry)
DE   Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW   Nucleotide analogue; alkaline phosphatase; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .501
FT                   /*tag=  a
FT                   /note= "std. IUPAC codes used"
PN   WO9524505-A.
PD   14-SEP-1995.
PF   07-MAR-1995; G00490.
PR   08-MAR-1994; EP-301636.
PA   (AMSH ) AMERSHAM INT PLC.
PI   Reeve MA,  Robinson PS;
DR   WPI; 95-328290/42.
PT   Modification of residual fluorescence labelled nucleotide analogues  - to
PT   prevent migration in electrophoretic sequencing gel and interference with
PT   base calling of DNA chains
PS   Example; Fig 1; 18pp; English.
CC   The invention concerns the modification of redisual fluorescence
CC   labelled nt analogues to prevent migration in electrophoretic
CC   sequencing gel and interference with base calling of DNA chains. The
CC   modification involves the use of a phosphatase enzyme to remove at
CC   least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC   dye-terminating kt. The template used was 1 microg. of M13 mp8. The
CC   primer was M13 Universal primer. Samples were recovered. Some were
CC   subjected to calf intestinal alkaline phospatase digestion. This
CC   includes T04866. T04867 was a control nt. which was not treated with
CC   alkaline phosphatase. The example shows that alkaline phosphatase
CC   treatment causes removal of the dye-terminator artefacts and allows
CC   for accurate base calling with the ABI analysis software.
SQ   Sequence   501 BP;    98 A;    136 C;    140 G;    121 T;


  Query Match             47.2%;  Score 69.8;  DB 1;  Length 501;
  Best Local Similarity   81.6%;  Pred. No. 1.7e-15;
  Matches   80;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      36 CGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACC 95
           |||    |   | ||||||||||||||||||||||||||||||||||||||||||||||
Db      44 CGCNCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACC 103

Qy      96 GTCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||||||||||||||       |   |||||||
Db     104 GTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 141


RESULT   4
T75006/c
ID   T75006 standard; DNA; 1091 BP.
AC   T75006;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 6.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS   Human retrovirus.
PN   WO9725431-A1.
PD   17-JUL-1997.
PF   10-JAN-1997; U00398.
PR   10-JAN-1996; US-587329.
PA   (CORI-) CORIXA CORP.
PI   Frudakis TN,  Smith JM;
DR   WPI; 97-384982/35.
PT   Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT   for the prognosis, diagnosis and monitoring of human cancers,
PT   especially breast cancer
PS   Claim 10; Page 31-32; 74pp; English.
CC   Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC   and 11-22 (T75003-10) were obtd. by screening human genomic
CC   libraries using human breast tumour-associated retroviral element
CC   B18Ag1 (see also T75002) as probe.  These non-contiguous sequences
CC   lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC   retrovirus genome (see also T75001).  B18Ag1 and the other
CC   retroviral sequences can be used in genetic vaccines and for the
CC   prognosis, diagnosis and monitoring of human breast cancer.
SQ   Sequence   1091 BP;    79 A;    350 C;    97 G;    248 T;


  Query Match             46.4%;  Score 68.6;  DB 1;  Length 1091;
  Best Local Similarity   80.8%;  Pred. No. 6e-15;
```

```
  Matches   80;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      17 ACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGA 76
           | |||| ||       ||  ||    |   | ||||||||||||||||||||||||||||
Db     169 ATTAGAGCTCGCGGCCGCGAGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGA 110

Qy      77 TATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           |||||||||||||||||||||||||||||||||||||||
Db     109 TATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 71


RESULT   5
T43794/c
ID   T43794 standard; DNA; 5356 BP.
AC   T43794;
DT   18-FEB-1997  (first entry)
DE   Plasmid pRIPHAT (rat insulin promoter-human IAPP transgene).
KW   Type II diabetes mellitus; transgenic animal model; pancreas;
KW   islet; beta cell; islet amyloid polypeptide; IAPP; hyperglycaemia;
KW   glycouria; diabetic glomerulosclerosis; plasmid pRIPHAT; ds;
KW   cyclic.
OS   Chimeric Homo sapiens;
OS   Chimeric Rattus sp.;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   promoter        719. .1594
FT                   /*tag= a
FT                   /label= RIP-II_promoter
FT                   /note= "rat insulin II promoter"
FT   cds             1589. .1866
FT                   /*tag= b
FT                   /label= IAPP
FT                   /note= "human IAPP coding sequence (Claim 7)"
FT   intron          1861. .2580
FT                   /*tag= c
FT                   /label= Albumin_intron
FT                   /note= "human albumin intron I"
FT   terminator      2575. .3119
FT                   /*tag= d
FT                   /label= Poly-A
FT                   /note= "human GAPDH gene poly-A and RNA
FT                   termination region"
PN   WO9637612-A1.
PD   28-NOV-1996.
PF   24-APR-1996; IB0371.
PR   23-MAY-1995; US-446935.
PA   (PFIZ ) PFIZER INC.
PI   Carty MD,  Kreutter DK,  Soeller WC;
DR   WPI; 97-021221/02.
PT   Recombinant DNA for expression of islet amyloid polypeptide  - to
PT   develop prods. for use in diagnosis, study and treatment of
PT   disorders, e.g. diabetes and obesity
PS   Claim 6; Page 21-23; 46pp; English.
CC   Plasmid pRIPHAT (T43794) allows pancreatic beta cell-specific
CC   expression of human islet amyloid-associated polypeptide (IAPP)
CC   under the regulation of the rat insulin II promoter in both cell
CC   lines and in transgenic animals (models of type II diabetes
CC   mellitus).  It was constructed by PCR amplification (see also
CC   T43795-804) of the relevant DNA fragments and their insertion into
CC   a pBluescript SK(-) framework.  The transgenic animals develop
CC   amyloid plaque deposits in islet cells, fasting hyperglycaemia,
CC   glycouria and diabetic glomerulosclerosis.  They can be screened
CC   for treatments that modulate disease progression.  Cell lines can
CC   be screened for treatments that inhibit human IAPP expression.
SQ   Sequence   5356 BP;    1368 A;    1294 C;    1274 G;    1420 T;


  Query Match             45.9%;  Score 68;  DB 1;  Length 5356;
  Best Local Similarity   88.1%;  Pred. No. 1.7e-14;
  Matches   74;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      32 TGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGA 91
           | |||  ||      ||| ||||||||||||||||||||||||||||||||||||||||||
Db     743 TCCACTTGGAGTGGTTGGGGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGA 684

Qy      92 TACCGTCGACCTCGAGGGGGGGCC 115
           ||||||||||||||||||||||||
Db     683 TACCGTCGACCTCGAGGGGGGGCC 660


RESULT   6
T75005/c
ID   T75005 standard; DNA; 1122 BP.
AC   T75005;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 3.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS   Human retrovirus.
PN   WO9725431-A1.
PD   17-JUL-1997.
PF   10-JAN-1997; U00398.
PR   10-JAN-1996; US-587329.
PA   (CORI-) CORIXA CORP.
PI   Frudakis TN,  Smith JM;
DR   WPI; 97-384982/35.
PT   Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT   for the prognosis, diagnosis and monitoring of human cancers,
PT   especially breast cancer
PS   Claim 10; Page 30-31; 74pp; English.
CC   Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC   and 11-22 (T75003-10) were obtd. by screening human genomic
CC   libraries using human breast tumour-associated retroviral element
CC   B18Ag1 (see also T75002) as probe.  These non-contiguous sequences
CC   lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC   retrovirus genome (see also T75001).  B18Ag1 and the other
CC   retroviral sequences can be used in genetic vaccines and for the
CC   prognosis, diagnosis and monitoring of human breast cancer.
SQ   Sequence   1122 BP;    260 A;    316 C;    185 G;    279 T;


  Query Match             45.7%;  Score 67.6;  DB 1;  Length 1122;
  Best Local Similarity   80.6%;  Pred. No. 1.4e-14;
  Matches   79;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      18 CTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGAT 77
           ||| |    |  ||| |  |     |   | |||||||||||||||||||||||||||||
Db     139 CTAAAATCCTCCATTTCTTGGTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGAT 80

Qy      78 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           |||||||||||||||||||||||||||||||||||||||
Db      79 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 42


RESULT   7
X20513
ID   X20513 standard: DNA; 651 BP.
AC   X20513;
DT   05-MAY-1999  (first entry)
DE   Polynucleotide sequence from the genome of Treponema pallidum.
KW   Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW   enzyme production; ds.
OS   Treponema pallidum.
PN   WO9859034-A2.
PD   30-DEC-1998.
PF   23-JUN-1998; U13041.
PR   24-JUN-1997; US-050667.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Fraser CM;
DR   WPI; 99-081273/07.
PT   New isolated Treponema pallidum nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of T. pallidum infections, particularly syphilis
PS   Claim 1; Page 257; 1150pp; English.
CC   X20500-21243 represent polynucleotide sequences from the genome of
CC   Treponema pallidum. The sequences can be used for detection,
CC   diagnosis, characterisation, prevention and therapy for T. pallidum
```

```
CC   infections, particularly syphilis. They can also be used for detecting
CC   diseases related to Borrelia infections in animals, and for the
CC   production of biosynthetic products such as enzymes.
SQ   Sequence   651 BP;    180 A;    167 C;    152 G;    149 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 651;
  Best Local Similarity   98.6%;  Pred. No. 1.3e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 135

Qy    107 GGGGGGGCC 115
          |||||||||
Db    136 GGGGGGGCC 144


RESULT   8
V31294/c
ID   V31294 standard; DNA; 752 BP.
AC   V31294;
DT   01-OCT-1998  (first entry)
DE   E. coli J96 pathogenicity island contig #108.
KW   PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW   PAI V; pheV; vaccine; protective immune response; ds.
OS   Escherichia coli.
PN   WO9822575-A2.
PD   28-MAY-1998.
PF   21-NOV-1997; U21347.
PR   14-OCT-1997; US-061953.
PR   22-NOV-1996; US-031626.
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYWI-) UNIV WISCONSIN.
PI   Choi GH, Dillon PJ, Welch RA;
DR   WPI; 98-312461/27.
PT   New isolated uropathogenic E. coli nucleotide sequences - used to
PT   develop products for the detection of pathogenic E. coli and to
PT   elicit an immune response to pathogenic E. coli
PS   Claim 21; Page 221-222; 250pp; English.
CC   This sequence represents a E. coli strain J96 contig containing
CC   pathogenicity island (PAI) sequences, and represents a nucleic acid
CC   molecule of the invention. PAIs are large fragments of DNA which comprise
CC   pathogenicity determinants. The sequences of the invention are taken from
CC   PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC   on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC   approximately 94 min (at pheR) on the E. coli chromosome and is
CC   approximately 160 kb in size. Antibodies specific to the proteins encoded
CC   by the PAI open reading frames of the invention can be used in kits to
CC   detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC   a protective immune response in an animal to the uropathogenic E. coli
CC   strain J96.
SQ   Sequence   752 BP;    162 A;    213 C;    203 G;    172 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 752;
  Best Local Similarity   98.6%;  Pred. No. 1.4e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 15

Qy    107 GGGGGGGCC 115
          |||||||||
Db     14 GGGGGGGCC 6


RESULT   9
Q48463/c
ID   Q48463 standard; DNA; 3792 BP.
AC   Q48463;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host4 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 9; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463). It
CC   is thermosensitive in all hosts tested, including E.coli, and must be
CC   maintained at 28 deg.C.
SQ   Sequence   3792 BP;    1249 A;    552 C;    742 G;    1249 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 3792;
  Best Local Similarity   98.6%;  Pred. No. 2.4e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3543 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3484

Qy    107 GGGGGGGCC 115
          |||||||||
Db   3483 GGGGGGGCC 3475


RESULT  10
T39485
ID   T39485 standard; DNA; 4226 BP.
AC   T39485;
DT   21-MAY-1997  (first entry)
DE   Human steroidogenesis acute regulatory protein genomic DNA.
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; ds.
OS   Homo sapiens.
PN   WO9629338-A1.
PD   26-SEP-1996.
PF   22-MAR-1996; U03896.
PR   23-MAR-1995; US-410540.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
PI   Lin D,  Miller WL,  Strauss JF;
DR   WPI; 96-443130/44.
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
PS   Claim 1; Pages 23-25; 89pp; English.
CC   The present sequence encodes the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
```

```
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc abd pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
SQ   Sequence   4226 BP;    940 A;    1132 C;    1144 G;    984 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 4226;
  Best Local Similarity   98.6%;  Pred. No. 2.5e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4147 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4206

Qy     107 GGGGGGGCC 115
           |||||||||
Db    4207 GGGGGGGCC 4215


RESULT  11
Q48464/c
ID   Q48464 standard; DNA; 5234 BP.
AC   Q48464;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host5 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    36. .1496
FT                   /*tag= a
FT                   /standard_name= ORI
FT                   /note= "origin of replication from pBR322"
FT   misc_feature    2640. .4383
FT                   /*tag= b
FT                   /note= "from pGK12 (derived from pWV01)"
FT   misc_feature    4384. .4786
FT                   /*tag= c
FT                   /note= "from pUB110"
FT   misc_feature    4787. .5234
FT                   /*tag= d
FT                   /note= "from pSK"
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 10; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the 1.4kb AvaI-AlwNI fragment of
CC   pBR322 (containing the origin of replication) was inserted into
CC   NsiI-cleaved pG+host4 to give pG+host5 (Q48464).The pBR322 ORI
CC   allows maintenance of the plasmid in E.coli at 37 deg.C. while the
CC   heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
CC   bacteria.
SQ   Sequence   5234 BP;    1569 A;    973 C;    1133 G;    1559 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 5234;
  Best Local Similarity   98.6%;  Pred. No. 2.7e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4985 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4926

Qy     107 GGGGGGGCC 115
           |||||||||
Db    4925 GGGGGGGCC 4917


RESULT  12
Q48465/c
ID   Q48465 standard; DNA; 6722 BP.
AC   Q48465;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host6 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Disclosure; Fig 11; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the AvaI-EcoRI fragment of pBR322
CC   (containing the origin of replication and the ampicillin resistance
CC   gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
CC   pG+host6 (Q48466).The pBR322 ORI allows maintenance of the plasmid in
CC   E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC   at 28 deg.C in gram-positive bacteria.
SQ   Sequence   6722 BP;    1961 A;    1278 C;    1503 G;    1980 T;


  Query Match             45.5%;  Score 67.4;  DB 1;  Length 6722;
  Best Local Similarity   98.6%;  Pred. No. 3e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6473 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 6414

Qy     107 GGGGGGGCC 115
           |||||||||
Db    6413 GGGGGGGCC 6405


RESULT  13
X24730
ID   X24730 standard; DNA; 12814 BP.
AC   X24730;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP targetting vector pMTI-2398.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
```

```
KW   animal model; Alzheimer's disease; vector; pMTI-2398;
KW   Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1932. .2276
FT                   /*tag= a
FT                   /standard_name= Swedish-FAD APP
FT                   /note= "encodes W97997"
FT   mat_peptide     5369. .6160
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97997, W97998.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 96-103; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP targetting
CC   vector pMTI-2398.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5007,
CC   created using vector pMTI-2398, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced.  The targetted Swedish-FAD
CC   m/hAPP gene expressed m/hAPP protein at levels approaching those
CC   observed for mouse APP in brain.
SQ   Sequence   12814 BP;    1865 A;    1850 C;    1962 G;    1939 T;

  Query Match             45.5%;  Score 67.4;  DB 1;  Length 12814;
  Best Local Similarity   98.6%;  Pred. No. 3.7e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9886 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 9945

Qy     107 GGGGGGGCC 115
           |||||||||
Db    9946 GGGGGGGCC 9954


RESULT  14
X24731
ID   X24731 standard; DNA; 15692 BP.
AC   X24731;
DT   21-JUN-1999  (first entry)
DE   London-FAD APP targetting vector pMTI-2453.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2453;
KW   London-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /standard_name= "Swedish-FAD"
FT                   /note= "encodes W97999"
FT   mutation        replace(4990,"")
FT                   /*tag= b
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= c
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W97999.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 104-113; 209pp; English.
CC   This is the nucleotide sequence of London-FAD APP targetting
CC   vector pMTI-2453.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5401,
CC   created using vector pMTI-2453, the London mutation. i.e.
CC   V(717)L, was also introduced.
SQ   Sequence   15692 BP;    2423 A;    2730 C;    2786 G;    2541 T;

  Query Match             45.5%;  Score 67.4;  DB 1;  Length 15692;
  Best Local Similarity   98.6%;  Pred. No. 4e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12764 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 12823

Qy     107 GGGGGGGCC 115
           |||||||||
Db   12824 GGGGGGGCC 12832


RESULT  15
X24732
ID   X24732 standard; DNA; 15692 BP.
AC   X24732;
DT   21-JUN-1999  (first entry)
DE   Swedish/London-FAD APP targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2454;
KW   London-FAD APP; Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
```

```
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /note= "encodes W98000"
FT   mutation        replace(4849,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4989,"")
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98000.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 114-123; 209pp; English.
CC   This is the nucleotide sequence of Swedish/London-FAD APP targetting
CC   vector pMTI-2454. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5103,
CC   created using vector pMTI-2454, the London mutation. i.e.
CC   V(717)L, and the Swedish mutation, i.e. KM(670,671)NL, were
CC   introduced. The targetted Swedish/London-FAD m/hAPP gene expressed
CC   m/hAPP protein at levels approaching those observed for mouse APP
CC   in brain.
SQ   Sequence   15692 BP;    2423 A;    2731 C;    2783 G;    2542 T;


  Query Match            45.5%;  Score 67.4;  DB 1;  Length 15692;
  Best Local Similarity  98.6%;  Pred. No. 4e-14;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12764 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 12823

Qy    107 GGGGGGGCC 115
          |||||||||
Db  12824 GGGGGGGCC 12832


Search completed: May 29, 2000, 21:58:27
Job time: 38175 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:00 ; Search time 621.83 Seconds
(without alignments)
30.937 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	68.4	46.2	793	3	US-08-967-101-121	Sequence 121, App
c	2	68.4	46.2	793	4	US-08-592-541-121	Sequence 121, App
c	3	67.4	45.5	3792	4	US-08-992-334-1	Sequence 1, Appli
c	4	67.4	45.5	3792	5	US-08-302-752-1	Sequence 1, Appli
	5	67.4	45.5	4016	2	US-08-410-540-3	Sequence 3, Appli
c	6	67.4	45.5	5234	4	US-08-992-334-2	Sequence 2, Appli
c	7	67.4	45.5	5234	5	US-08-302-752-2	Sequence 2, Appli
c	8	67.4	45.5	6722	4	US-08-992-334-3	Sequence 3, Appli
c	9	67.4	45.5	6722	5	US-08-302-752-3	Sequence 3, Appli
	10	65.4	44.2	9318	3	US-08-793-610-6	Sequence 6, Appli
	11	62.8	42.4	685	2	US-08-463-115-56	Sequence 56, Appl
	12	62.8	42.4	685	2	US-08-465-388-56	Sequence 56, Appl
	13	59.6	40.3	4164	1	US-08-204-675-1	Sequence 1, Appli
	14	59.6	40.3	4164	3	US-08-660-754-1	Sequence 1, Appli
	15	59.6	40.3	4164	4	US-08-796-364-1	Sequence 1, Appli
	16	59.6	40.3	4164	6	PCT-US95-02520-1	Sequence 1, Appli
	17	58.4	39.5	5178	3	US-08-474-169-2	Sequence 2, Appli
	18	56	37.8	88	1	US-08-144-602B-15	Sequence 15, Appl
c	19	56	37.8	1023	1	US-08-198-446B-1	Sequence 1, Appli
c	20	56	37.8	1023	3	US-08-870-693-1	Sequence 1, Appli
c	21	56	37.8	2150	1	US-08-198-446B-10	Sequence 10, Appl
c	22	56	37.8	2150	3	US-08-870-693-10	Sequence 10, Appl
c	23	55	37.2	834	3	US-08-967-101-113	Sequence 113, App
c	24	55	37.2	834	4	US-08-592-541-113	Sequence 113, App
	25	51.6	34.9	6505	3	US-08-793-610-5	Sequence 5, Appli
	26	51	34.5	3341	3	US-08-868-577-18	Sequence 18, Appl
c	27	50	33.8	780	3	US-08-967-101-116	Sequence 116, App
c	28	50	33.8	780	4	US-08-592-541-116	Sequence 116, App
c	29	49.8	33.6	1200	1	US-08-011-398B-3	Sequence 3, Appli
c	30	49.8	33.6	1200	2	US-08-464-051-3	Sequence 3, Appli
c	31	49.8	33.6	1200	3	US-08-462-498-3	Sequence 3, Appli
c	32	49.8	33.6	6244	1	US-08-076-726-15	Sequence 15, Appl
c	33	49.8	33.6	6244	1	US-08-260-452-8	Sequence 8, Appli
c	34	49.8	33.6	6244	3	US-08-481-970-8	Sequence 8, Appli
c	35	49.8	33.6	6244	4	US-08-897-719-8	Sequence 8, Appli
c	36	48.2	32.6	2308	1	US-08-325-071-62	Sequence 62, Appl
	37	48	32.4	1612	1	US-08-343-733A-1	Sequence 1, Appli
	38	48	32.4	2764	4	US-08-465-971B-1	Sequence 1, Appli
c	39	47.2	31.9	2185	4	US-08-467-948A-3	Sequence 3, Appli
c	40	47	31.8	1249	2	US-08-463-115-35	Sequence 35, Appl
c	41	47	31.8	1249	2	US-08-465-388-35	Sequence 35, Appl
c	42	47	31.8	1997	4	US-08-667-809B-3	Sequence 3, Appli
c	43	47	31.8	2126	3	US-08-789-354-1	Sequence 1, Appli
	44	47	31.8	2605	3	US-08-680-395-4	Sequence 4, Appli
	45	47	31.8	5534	2	US-08-452-267-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT   1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-967-101-121


  Query Match             46.2%;  Score 68.4;  DB 3;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 3.8e-14;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;
```

```
Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   2
US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-592-541-121


  Query Match             46.2%;  Score 68.4;  DB 4;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 3.8e-14;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   3
US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
;    IMMEDIATE SOURCE:
;      CLONE:  pG+host4
US-08-992-334-1


  Query Match             45.5%;  Score 67.4;  DB 4;  Length 3792;
  Best Local Similarity   98.6%;  Pred. No. 1.3e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3543 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3484

Qy     107 GGGGGGGCC 115
           |||||||||
Db    3483 GGGGGGGCC 3475


RESULT   4
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-1


  Query Match             45.5%;  Score 67.4;  DB 5;  Length 3792;
  Best Local Similarity   98.6%;  Pred. No. 1.3e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3543 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3484

Qy     107 GGGGGGGCC 115
           |||||||||
Db    3483 GGGGGGGCC 3475


RESULT   5
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F.
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,540
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-238/00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4016 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1098..1283
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1620..1733
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2047..2174
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2267..2425
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2567..2751
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2828..2921
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3031..3765
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1433..1434
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2208..2209
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2781..2782
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2956..2957
;      OTHER INFORMATION:  /note= "interruption of sequence data"
US-08-410-540-3


  Query Match             45.5%;  Score 67.4;  DB 2;  Length 4016;
  Best Local Similarity   98.6%;  Pred. No. 1.3e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3937 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3996

Qy     107 GGGGGGGCC 115
           |||||||||
Db    3997 GGGGGGGCC 4005


RESULT   6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
```

```
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-2


  Query Match             45.5%;  Score 67.4;  DB 4;  Length 5234;
  Best Local Similarity   98.6%;  Pred. No. 1.4e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4985 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4926

Qy     107 GGGGGGGCC 115
           |||||||||
Db    4925 GGGGGGGCC 4917


RESULT   7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-2


  Query Match             45.5%;  Score 67.4;  DB 5;  Length 5234;
  Best Local Similarity   98.6%;  Pred. No. 1.4e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4985 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4926

Qy     107 GGGGGGGCC 115
           |||||||||
Db    4925 GGGGGGGCC 4917


RESULT   8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-3
```

```
  Query Match             45.5%;  Score 67.4;  DB 4;  Length 6722;
  Best Local Similarity   98.6%;  Pred. No. 1.5e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6473 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 6414

Qy     107 GGGGGGGCC 115
           |||||||||
Db    6413 GGGGGGGCC 6405


RESULT   9
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-3


  Query Match             45.5%;  Score 67.4;  DB 5;  Length 6722;
  Best Local Similarity   98.6%;  Pred. No. 1.5e-13;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6473 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 6414

Qy     107 GGGGGGGCC 115
           |||||||||
Db    6413 GGGGGGGCC 6405


RESULT  10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6


  Query Match             44.2%;  Score 65.4;  DB 3;  Length 9318;
  Best Local Similarity   98.5%;  Pred. No. 7.6e-13;
  Matches   66;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      49 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 108
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5654 GGGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 5713

Qy     109 GGGGGCC 115
           |||||||
Db    5714 GGGGGCC 5720


RESULT  11
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
```

```
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,115
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:           Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5703221ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/301
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-463-115-56


  Query Match             42.4%;  Score 62.8;  DB 2;  Length 685;
  Best Local Similarity   92.8%;  Pred. No. 2.7e-12;
  Matches   64;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||
Db     148 TAGTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGA 207

Qy     107 GGGGGGGCC 115
           ||||||| |
Db     208 GGGGGGGNC 216


RESULT  12
US-08-465-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,388
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:           Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5753488ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/300
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-465-388-56


  Query Match             42.4%;  Score 62.8;  DB 2;  Length 685;
  Best Local Similarity   92.8%;  Pred. No. 2.7e-12;
  Matches   64;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||
Db     148 TAGTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGA 207

Qy     107 GGGGGGGCC 115
           ||||||| |
Db     208 GGGGGGGNC 216


RESULT  13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 5677170
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,675
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-204-675-1


  Query Match             40.3%;  Score 59.6;  DB 1;  Length 4164;
  Best Local Similarity   75.5%;  Pred. No. 5.2e-11;
  Matches   74;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      10 TTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGG 69
           | ||| |  |  | |    |   |  |||    |   | ||||||||||||||||||||
Db    2005 TCGAACATGTTCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGG 2064

Qy      70 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           |||||||||||||||||||||||||||||||||||||
Db    2065 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  14
US-08-660-754-1
; Sequence 1, Application US/08660754
; Patent No. 5843772
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/660,754
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-660-754-1


  Query Match             40.3%;  Score 59.6;  DB 3;  Length 4164;
  Best Local Similarity   75.5%;  Pred. No. 5.2e-11;
  Matches   74;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      10 TTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGG 69
           | ||| |  |  | |    |   |  |||    |   | ||||||||||||||||||||
Db    2005 TCGAACATGTTCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGG 2064

Qy      70 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           |||||||||||||||||||||||||||||||||||||
Db    2065 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  15
US-08-796-364-1
; Sequence 1, Application US/08796364
; Patent No. 5968785
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/796,364
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
```

; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: pAT-1
US-08-796-364-1

Query Match 40.3%; Score 59.6; DB 4; Length 4164;
Best Local Similarity 75.5%; Pred. No. 5.2e-11;
Matches 74; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy      10 TTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGG 69
           | ||| |  |  | |    |   |  |||    |   |||||||||||||||||||||
Db    2005 TCGAACATGTTCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGG 2064

Qy      70 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           |||||||||||||||||||||||||||||||||||||
Db    2065 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102
```

Search completed: May 29, 2000, 22:09:05
Job time: 38746 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:30 ; Search time 2276.24 Seconds
(without alignments)
290.102 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	148	28	US-08-935-377-7	Sequence 7, Appli
2	137	92.6	149	28	US-08-935-377-8	Sequence 8, Appli

	3	136	91.9	150	28	US-08-935-377-9 Sequence 9, Appli
	4	132	89.2	145	28	US-08-935-377-6 Sequence 6, Appli
	5	70.6	47.7	858	40	US-09-123-912-91 Sequence 91, Appl
	6	70.6	47.7	858	43	US-09-221-107-91 Sequence 91, Appl
	7	70.6	47.7	858	46	US-09-285-479-91 Sequence 91, Appl
	8	69.4	46.9	224	13	US-08-377-383-4 Sequence 4, Appli
	9	69.4	46.9	224	18	US-08-693-573-4 Sequence 4, Appli
c	10	68.4	46.2	793	15	US-08-431-048A-121 Sequence 121, App
c	11	68.4	46.2	793	16	US-08-496-841-121 Sequence 121, App
c	12	68.4	46.2	793	17	US-08-509-359-121 Sequence 121, App
c	13	68.4	46.2	793	17	US-08-509-359B-121 Sequence 121, App
c	14	68.4	46.2	793	40	US-09-124-523-121 Sequence 121, App
c	15	68.4	46.2	793	40	US-09-124-698-121 Sequence 121, App
c	16	68.4	46.2	793	40	US-09-127-480-121 Sequence 121, App
c	17	68	45.9	5356	15	US-08-446-935-1 Sequence 1, Appli
c	18	67.4	45.5	147	38	US-09-072-433-31 Sequence 31, Appl
c	19	67.4	45.5	157	63	US-60-029-960-595 Sequence 595, App
	20	67.4	45.5	245	54	US-09-411-999-19130 Sequence 19130, A
	21	67.4	45.5	381	49	US-09-332-782-8570 Sequence 8570, Ap
	22	67.4	45.5	381	57	US-09-515-694-8570 Sequence 8570, Ap
	23	67.4	45.5	393	54	US-09-411-999-13627 Sequence 13627, A
	24	67.4	45.5	405	49	US-09-332-782-8523 Sequence 8523, Ap
	25	67.4	45.5	405	57	US-09-515-694-8523 Sequence 8523, Ap
	26	67.4	45.5	420	49	US-09-332-782-8595 Sequence 8595, Ap
	27	67.4	45.5	420	57	US-09-515-694-8595 Sequence 8595, Ap
	28	67.4	45.5	422	49	US-09-332-782-9309 Sequence 9309, Ap
	29	67.4	45.5	422	57	US-09-515-694-9309 Sequence 9309, Ap
	30	67.4	45.5	423	54	US-09-411-999-34199 Sequence 34199, A
	31	67.4	45.5	433	42	US-09-179-473-2334 Sequence 2334, Ap
	32	67.4	45.5	433	49	US-09-328-351-2334 Sequence 2334, Ap
	33	67.4	45.5	442	42	US-09-179-473-2495 Sequence 2495, Ap
	34	67.4	45.5	442	49	US-09-328-351-2495 Sequence 2495, Ap
	35	67.4	45.5	448	42	US-09-179-473-2255 Sequence 2255, Ap
	36	67.4	45.5	448	49	US-09-328-351-2255 Sequence 2255, Ap
	37	67.4	45.5	454	42	US-09-179-473-2100 Sequence 2100, Ap
	38	67.4	45.5	454	42	US-09-179-473-2360 Sequence 2360, Ap
	39	67.4	45.5	454	49	US-09-328-351-2100 Sequence 2100, Ap
	40	67.4	45.5	454	49	US-09-328-351-2360 Sequence 2360, Ap
c	41	67.4	45.5	752	32	US-08-976-259-108 Sequence 108, App
c	42	67.4	45.5	752	64	US-60-031-626-108 Sequence 108, App
c	43	67.4	45.5	827	82	US-60-126-265-996 Sequence 996, App
c	44	67.4	45.5	2958	37	US-09-040-961-6 Sequence 6, Appli
c	45	67.4	45.5	2958	41	US-09-130-115-6 Sequence 6, Appli

ALIGNMENTS

```
RESULT   1
US-08-935-377-7
; Sequence 7, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  148 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-7


  Query Match           100.0%;  Score 148;  DB 28;  Length 148;
  Best Local Similarity  100.0%;  Pred. No. 1e-39;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCG 60

Qy      61 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 120

Qy     121 AACTAATTTTGTTTTTGTGGGCCCGGCC 148
           ||||||||||||||||||||||||||||
Db     121 AACTAATTTTGTTTTTGTGGGCCCGGCC 148


RESULT   2
US-08-935-377-8
; Sequence 8, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-8
```

```
  Query Match             92.6%;  Score 137;  DB 28;  Length 149;
  Best Local Similarity   99.3%;  Pred. No. 5e-36;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG-GTGGATCCCCC 59
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60

Qy     60 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120

Qy    120 TAACTAATTTTGTTTTTGTGGGCCCGGCC 148
          ||||||||||||||||||||||||||||
Db    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149


RESULT   3
US-08-935-377-9
; Sequence 9, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-9


  Query Match             91.9%;  Score 136;  DB 28;  Length 150;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-35;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG--GTGGATCCCC 58
          |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60

Qy     59 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 118
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120

Qy    119 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 148
          ||||||||||||||||||||||||||||||
Db    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150


RESULT   4
US-08-935-377-6
; Sequence 6, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-6


  Query Match             89.2%;  Score 132;  DB 28;  Length 145;
  Best Local Similarity   98.0%;  Pred. No. 2.4e-34;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCG 60
          |||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCC---GTGGATCCCCCG 57

Qy     61 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58 GGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAACT 117

Qy    121 AACTAATTTTGTTTTTGTGGGCCCGGCC 148
          ||||||||||||||||||||||||||||
Db    118 AACTAATTTTGTTTTTGTGGGCCCGGCC 145


RESULT   5
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE:  210121.455C1
;  CURRENT APPLICATION NUMBER: US/09/123,912A
;  CURRENT FILING DATE:  1998-07-27
;  PRIOR APPLICATION NUMBER:  09/040,802
;  PRIOR FILING DATE:  1998-03-18
```

```
;  NUMBER OF SEQ ID NOS:  114
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (570)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (591)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (655)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (664)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (667)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (683)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (711)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (792)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-91

  Query Match           47.7%;  Score 70.6;  DB 40;  Length 858;
  Best Local Similarity  84.9%;  Pred. No. 1.6e-13;
  Matches   79;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      23 TCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAA 82
           || |   ||| |||    | |   | ||||||||||||||||||||||||||||||||||

Db     331 tcaacgaattccaccacactggactagtggatcccccgggctgcaggaattcgatatcaa 390

Qy      83 GCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           |||||||||||||||||||||||||||||||||
Db     391 gcttatcgataccgtcgacctcgaggggggcc 423


RESULT   6
US-09-221-107-91
; Sequence 91, Application US/09221107
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE: 210121.455C2
;  CURRENT APPLICATION NUMBER: US/09/221,107
;  CURRENT FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 161
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (570)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (591)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (655)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (664)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (667)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (683)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (711)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (792)
```

```
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-91


  Query Match             47.7%;  Score 70.6;  DB 43;  Length 858;
  Best Local Similarity   84.9%;  Pred. No. 1.6e-13;
  Matches   79;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      23 TCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAA 82
           || |   ||| |||    | |   | ||||||||||||||||||||||||||||||||||
Db     331 tcaacgaattccaccacactggactagtggatccccgggctgcaggaattcgatatcaa 390

Qy      83 GCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           |||||||||||||||||||||||||||||||||
Db     391 gcttatcgataccgtcgacctcgagggggggcc 423


RESULT   7
US-09-285-479-91
; Sequence 91, Application US/09285479
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: LUNG CANCER
;  FILE REFERENCE: 210121.455C3
;  CURRENT APPLICATION NUMBER: US/09/285,479
;  CURRENT FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 172
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(858)
;   OTHER INFORMATION: n = A,T,C or G
US-09-285-479-91


  Query Match             47.7%;  Score 70.6;  DB 46;  Length 858;
  Best Local Similarity   84.9%;  Pred. No. 1.6e-13;
  Matches   79;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      23 TCTATTTATTGCACGCGGCCGCCATGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAA 82
           || |   ||| |||    | |   | ||||||||||||||||||||||||||||||||||
Db     331 tcaacgaattccaccacactggactagtggatccccgggctgcaggaattcgatatcaa 390

Qy      83 GCTTATCGATACCGTCGACCTCGAGGGGGGGCC 115
           |||||||||||||||||||||||||||||||||
Db     391 gcttatcgataccgtcgacctcgagggggggcc 423


RESULT   8
US-08-377-383-4
; Sequence 4, Application US/08377383
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru, FUKAMI,Tadashi, HIYOSHI,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of
;    TITLE OF INVENTION:  Nucleic Acid
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman Darby & Cushman
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskett, 3.5 inch, 720Kb storage
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1 or ASCII editors
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/377,383
;      FILING DATE:  25-JAN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6971/1994
;      FILING DATE:  26-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glenn J. Perry
;      REGISTRATION NUMBER:  28458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-377-383-4


  Query Match             46.9%;  Score 69.4;  DB 13;  Length 224;
  Best Local Similarity   87.4%;  Pred. No. 2.9e-13;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 145

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     146 GGGGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT   9
US-08-693-573-4
; Sequence 4, Application US/08693573
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru,
;    APPLICANT:  FUKAMI,Tadashi,
;    APPLICANT:  KOJIMA,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of                Nucl
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY CUSHMAN
;      ADDRESSEE:  INTELLECTUAL PROPERTY GROUP OF
;      ADDRESSEE:  PILLSBURY MADISON & SUTRO, L.L.P.
```

```
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 720Kb storage
;      COMPUTER:  IBM PC/XT/AT compatibles
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1 or ASCII editors
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,573
;      FILING DATE:  07-AUG-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/377,383
;      FILING DATE:  25-JAN-1995
;      APPLICATION NUMBER:  JP6971/1994
;      FILING DATE:  26-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas G. Wiseman
;      REGISTRATION NUMBER:  35046
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-693-573-4


  Query Match             46.9%;  Score 69.4;  DB 18;  Length 224;
  Best Local Similarity   87.4%;  Pred. No. 2.9e-13;
  Matches   76;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 145

Qy     107 GGGGGGGCCTAACTAACTAATTTTGTT 133
           |||||||||       |   |||||||
Db     146 GGGGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT  10
US-08-431-048A-121/c
; Sequence 121, Application US/08431048A
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  132
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Administrator, TESTA, HURWITZ &
;      ADDRESSEE:  THIBEAULT, LLP
;      STREET:  125 High Street
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,048A
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38,349
;      REFERENCE/DOCKET NUMBER:  CAN-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-431-048A-121


  Query Match             46.2%;  Score 68.4;  DB 15;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||        |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  11
US-08-496-841-121/c
; Sequence 121, Application US/08496841
;  GENERAL INFORMATION:
;    APPLICANT:  St. George-Hyslop, Peter
;    APPLICANT:  Rommens, Johanna M.
;    APPLICANT:  Fraser, Paul E.
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/496,841
;      FILING DATE:  28-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-496-841-121
```

```
  Query Match             46.2%;  Score 68.4;  DB 16;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  12
US-08-509-359-121/c
; Sequence 121, Application US/08509359
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359
;      FILING DATE:  31-JUL-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359-121


  Query Match             46.2%;  Score 68.4;  DB 17;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  13
US-08-509-359B-121/c
; Sequence 121, Application US/08509359B
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
;      STREET:  600 South Avenue West
;      CITY:  Westfield
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07090
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word, Version 6.0c
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359B
;      FILING DATE:  31-JUL-95
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Palisi, Thomas M.
;      REGISTRATION NUMBER:  36629
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 654-5000
;      TELEFAX:  (908) 654-7866
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359B-121


  Query Match             46.2%;  Score 68.4;  DB 17;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  14
US-09-124-523-121/c
; Sequence 121, Application US/09124523
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-523-121


  Query Match             46.2%;  Score 68.4;  DB 40;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  15
US-09-124-698-121/c
; Sequence 121, Application US/09124698
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,698
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-698-121


  Query Match             46.2%;  Score 68.4;  DB 40;  Length 793;
  Best Local Similarity   87.2%;  Pred. No. 8.8e-13;
  Matches   75;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      48 GGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

Qy     108 GGGGGGCCTAACTAACTAATTTTGTT 133
           ||||||||       |   |||||||
Db      59 GGGGGGCCCGGTACCCAGCTTTTGTT 34


Search completed:  May 30, 2000, 09:48:31
Job time:  60181 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:15 ; Search time 2192.43 Seconds
(without alignments)
273.613 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	70.4	47.6	682	96	AQ074693	AQ074693 CIT-HSP-2
c 2	68.4	46.2	608	96	AQ009167	AQ009167 CIT-HFP-2
c 3	67.8	45.8	528	84	B69688	B69688 CIT978SK-A-
4	67.4	45.5	108	84	B54292	B54292 CIT-HSP-201
5	67.4	45.5	141	44	AU037147	AU037147 AU037147
c 6	67.4	45.5	141	96	B95279	B95279 CIT-HSP-217
7	67.4	45.5	144	84	B80294	B80294 CIT-HSP-204
c 8	67.4	45.5	169	96	AQ013255	AQ013255 CIT-HSP-2
9	67.4	45.5	175	96	AQ041198	AQ041198 CIT-HSP-2
10	67.4	45.5	259	20	T48593	T48593 ph6f4_19/1T
11	67.4	45.5	275	27	AA052885	AA052885 T3551 Blo
c 12	67.4	45.5	386	84	B74162	B74162 CIT-HSP-202
c 13	67.4	45.5	462	96	AQ014512	AQ014512 CIT-HSP-2
c 14	67.4	45.5	519	84	B65534	B65534 CIT-HSP-202
15	67.4	45.5	547	96	AQ012629	AQ012629 CIT-HSP-2
16	67.4	45.5	552	96	AQ012216	AQ012216 CIT-HSP-2
17	67.4	45.5	575	96	AQ012221	AQ012221 CIT-HSP-2
c 18	67.4	45.5	596	96	AQ009165	AQ009165 CIT-HFP-2
c 19	67.4	45.5	607	84	B64363	B64363 CIT-HSP-202
20	67.4	45.5	632	84	B70299	B70299 CIT-HSP-202
c 21	67.4	45.5	638	96	AQ016737	AQ016737 CIT-HSP-2
22	67.4	45.5	647	84	B68775	B68775 CIT-HSP-205
23	67.4	45.5	666	84	B80296	B80296 CIT-HSP-204
24	67.4	45.5	700	96	AQ009153	AQ009153 CIT-HSP-2
c 25	67.2	45.4	576	38	AA752172	AA752172 96BS0243
c 26	66.4	44.9	541	96	AQ074695	AQ074695 CIT-HSP-2
27	65.8	44.5	513	96	AQ008437	AQ008437 CIT-HSP-2
c 28	65.8	44.5	541	96	AQ014519	AQ014519 CIT-HSP-2
c 29	63.2	42.7	134	84	B72088	B72088 CIT978SK-13
c 30	62.6	42.3	133	84	B26269	B26269 F10B5TF IGF
c 31	62.6	42.3	233	96	AQ012365	AQ012365 CIT-HSP-2
c 32	62.4	42.2	294	20	T26395	T26395 AB161A4F In
33	60	40.5	100	43	AI239372	AI239372 SWOvAFCAP
34	59	39.9	298	82	RPU68746	U68746 Riftia pach
c 35	58.2	39.3	630	96	AQ003238	AQ003238 RPCI11-19
c 36	57.8	39.1	145	84	B53695	B53695 CIT-HSP-201
c 37	57.4	38.8	603	38	AA752115	AA752115 96BS0066
38	56.8	38.4	97	41	AI058116	AI058116 SWAMCAC15
c 39	56.4	38.1	144	84	B65391	B65391 CIT-HSP-202
40	56	37.8	471	96	AQ024018	AQ024018 CpGR0359A
41	56	37.8	499	96	AQ023564	AQ023564 CpGR0482A
42	56	37.8	547	96	AQ023548	AQ023548 CpGR0051A
43	56	37.8	626	96	AQ024077	AQ024077 CpGR0415A
c 44	55.8	37.7	149	96	AQ044029	AQ044029 CIT-HSP-2
c 45	55.4	37.4	502	96	AQ023674	AQ023674 CpGR0007B

ALIGNMENTS

```
RESULT  1
AQ074693/c
LOCUS       AQ074693      682 bp    DNA             GSS       20-AUG-1998
DEFINITION  CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23,
            genomic survey sequence.
ACCESSION   AQ074693
VERSION     AQ074693.1  GI:3436811
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2301L23"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      156 a    187 c    187 g    151 t      1 others
ORIGIN
```

Query Match 47.6%; Score 70.4; DB 96; Length 682;
Best Local Similarity 92.5%; Pred. No. 2e-13;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 21

Qy     107 GGGGGGGCCTAACTAACTAA 126
           |||||||||  | || | ||
Db      20 GGGGGGGCCCGAGTACCCAA 1
```

```
RESULT  2
AQ009167/c
LOCUS       AQ009167      608 bp    DNA             GSS       27-JUN-1998
DEFINITION  CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13,
            genomic survey sequence.
ACCESSION   AQ009167
VERSION     AQ009167.1  GI:3128599
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
            CIT-HSP-2281N13.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .608
```

```
                     /organism="Homo sapiens"
                     /db_xref="GDB:7147360"
                     /db_xref="taxon:9606"
                     /clone="2281N13"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      141 a    164 c    166 g    137 t
ORIGIN


  Query Match             46.2%;  Score 68.4;  DB 96;  Length 608;
  Best Local Similarity   98.6%;  Pred. No. 9.3e-13;
  Matches   69;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 25

Qy     107 GGGGGGGCCT 116
           ||||||||||
Db      24 GGGGGGGCCT 15


RESULT   3
B69688/c
LOCUS       B69688         528 bp    DNA             GSS       18-JUN-1998
DEFINITION  CIT978SK-A-448E10.TVC CIT978SK Homo sapiens genomic clone A-448E10,
            genomic survey sequence.
ACCESSION   B69688
VERSION     B69688.1  GI:2708912
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /db_xref="GDB:5299607"
                     /db_xref="taxon:9606"
                     /clone="A-448E10"
                     /clone_lib="CIT978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
BASE COUNT      118 a    138 c    147 g    125 t
ORIGIN


  Query Match             45.8%;  Score 67.8;  DB 84;  Length 528;
  Best Local Similarity   97.2%;  Pred. No. 1.4e-12;
  Matches   69;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 18

Qy     107 GGGGGGGCCTA 117
           |||||||||| |
Db      17 GGGGGGGCCCA 7


RESULT   4
B54292
LOCUS       B54292         108 bp    DNA             GSS       20-JUN-1998
DEFINITION  CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
            genomic survey sequence.
ACCESSION   B54292
VERSION     B54292.1  GI:2608626
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7044003"
                     /db_xref="taxon:9606"
                     /clone="2017M1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       20 a     35 c     34 g     19 t
ORIGIN


  Query Match             45.5%;  Score 67.4;  DB 84;  Length 108;
  Best Local Similarity   98.6%;  Pred. No. 1.5e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 91

Qy     107 GGGGGGGCC 115
           |||||||||
Db      92 GGGGGGGCC 100


RESULT   5
AU037147
```

```
LOCUS       AU037147      141 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSB532, mRNA sequence.
ACCESSION   AU037147
VERSION     AU037147.1  GI:3983900
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2153012.
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSB532"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       26 a     40 c     42 g     28 t      5 others
ORIGIN
```

```
  Query Match             45.5%;  Score 67.4;  DB 44;  Length 141;
  Best Local Similarity   98.6%;  Pred. No. 1.6e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 126

Qy     107 GGGGGGGCC 115
           |||||||||
Db     127 GGGGGGGCC 135
```

```
RESULT   6
B95279/c
LOCUS       B95279        141 bp    DNA             GSS       26-JUN-1998
DEFINITION  CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
            genomic survey sequence.
ACCESSION   B95279
VERSION     B95279.1  GI:2977616
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Homo sapiens"
                     /db_xref="GDB:7103988"
                     /db_xref="taxon:9606"
                     /clone="2172N2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       24 a     41 c     41 g     35 t
ORIGIN
```

```
  Query Match             45.5%;  Score 67.4;  DB 96;  Length 141;
  Best Local Similarity   98.6%;  Pred. No. 1.6e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 16

Qy     107 GGGGGGGCC 115
           |||||||||
Db      15 GGGGGGGCC 7
```

```
RESULT   7
B80294
LOCUS       B80294        144 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
            genomic survey sequence.
ACCESSION   B80294
VERSION     B80294.1  GI:2867317
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2045D19.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="Homo sapiens"
                     /db_xref="GDB:7054558"
                     /db_xref="taxon:9606"
```

```
                     /clone="2045D19"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       34 a     43 c     41 g     26 t
ORIGIN


  Query Match           45.5%;  Score 67.4;  DB 84;  Length 144;
  Best Local Similarity   98.6%;  Pred. No. 1.6e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 127

Qy    107 GGGGGGGCC 115
          |||||||||
Db    128 GGGGGGGCC 136


RESULT   8
AQ013255/c
LOCUS       AQ013255      169 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
            genomic survey sequence.
ACCESSION   AQ013255
VERSION     AQ013255.1  GI:3185820
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2299C22.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .169
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154018"
                     /db_xref="taxon:9606"
                     /clone="2299C22"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       32 a     47 c     48 g     42 t
ORIGIN


  Query Match           45.5%;  Score 67.4;  DB 96;  Length 169;
  Best Local Similarity   98.6%;  Pred. No. 1.7e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 44

Qy    107 GGGGGGGCC 115
          |||||||||
Db     43 GGGGGGGCC 35


RESULT   9
AQ041198
LOCUS       AQ041198      175 bp    DNA             GSS       14-JUL-1998
DEFINITION  CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
            genomic survey sequence.
ACCESSION   AQ041198
VERSION     AQ041198.1  GI:3310469
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .175
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2335D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       43 a     50 c     47 g     35 t
ORIGIN


  Query Match           45.5%;  Score 67.4;  DB 96;  Length 175;
  Best Local Similarity   98.6%;  Pred. No. 1.7e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 130

Qy    107 GGGGGGGCC 115
          |||||||||
Db    131 GGGGGGGCC 139


RESULT  10
T48593
LOCUS       T48593        259 bp    mRNA            EST       02-FEB-1995
DEFINITION  ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph5f4_19/1TV, mRNA sequence.
```

```
ACCESSION   T48593
VERSION     T48593.1  GI:642793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L.,
            Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
            Sverdlov,E.D.
  TITLE     Outward Alu-primed hncDNA library
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Sverdlov ED
            Structure and Function of Human Genes
            Shemyakin Institute of Bioorganic Chemistry
            16/10 Miklukho-Maklaya, Moscow, 117871, Russia
            Tel: 70953306529
            Fax: 70953306538
            Email: sverd@humgen.siobc.msk.su.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ph6f4_19/1TV"
                     /clone_lib="Outward Alu-primed hncDNA library"
                     /note="Vector: pGEM-3Z; Site_1: EcoR1; Site_2: BamH1; The
                     library was constructed as described in [Obradovic, D.,
                     Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
                     Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
                     Sverdlov, E.D. (1993) Bioorganicheskaya khimia, 20,
                     919-930]. This protocol is based on nested primer strategy
                     using Alu- specific primers (ALN3 and TC-65) that direct
                     the hncDNA synthesis outward of Alu repeats."
BASE COUNT       62 a     66 c     72 g     59 t
ORIGIN

  Query Match             45.5%;  Score 67.4;  DB 20;  Length 259;
  Best Local Similarity   98.6%;  Pred. No. 1.8e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 113

Qy     107 GGGGGGGCC 115
           |||||||||
Db     114 GGGGGGGCC 122


RESULT  11
AA052885
LOCUS       AA052885      275 bp    mRNA            EST       13-SEP-1996
DEFINITION  T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
            brucei cDNA 5', mRNA sequence.
ACCESSION   AA052885
VERSION     AA052885.1  GI:1543898
KEYWORDS    EST.
SOURCE      Trypanosoma brucei brucei.
  ORGANISM  Trypanosoma brucei brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Osanya,A., Murphy,N.B. and Pelle,R.
  TITLE     Trypanosoma brucei brucei cDNAs
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:637578.
            Contact: Osanya A
            Unit 3
            International Livestock Research Institute
            Box 30709, Nairobi, Kenya
            Tel: 254 2 630 743
            Fax: 254 2 631 499
            Email: a.osanya@cgnet.com
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Trypanosoma brucei brucei"
                     /strain="Clone A4"
                     /db_xref="taxon:5702"
                     /clone_lib="Bloodstream form of serodeme ILTat1.1"
                     /note="cDNAs were generated from poly (A+) enriched mRNA
                     prepared from different developmental stages of T.b.brucei
                     by reverse transcription followed by PCR amplification
                     using mini-exon and oligo(dT) primers.  The cDNA generated
                     were utilized in RADES-PCR coupled with differential
                     hybridisations to identify differentially expressed mRNA
                     transcripts. The products which showed to be
                     differentially expressed were cloned pGEM -T vector.
                     These differentially expressed mRNA transcripts were (are
                     being) sequenced to generate differentially expressed
                     sequence tags."
BASE COUNT       61 a     84 c     74 g     56 t
ORIGIN

  Query Match             45.5%;  Score 67.4;  DB 27;  Length 275;
  Best Local Similarity   98.6%;  Pred. No. 1.8e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106

Qy     107 GGGGGGGCC 115
           |||||||||
Db     107 GGGGGGGCC 115


RESULT  12
B74162/c
LOCUS       B74162        386 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
            genomic survey sequence.
ACCESSION   B74162
VERSION     B74162.1  GI:2769849
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 386)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2028N1.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .386
                     /organism="Homo sapiens"
                     /db_xref="GDB:7048251"
```

```
                     /db_xref="taxon:9606"
                     /clone="2028N1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       81 a     98 c    109 g     98 t
ORIGIN


  Query Match             45.5%;  Score 67.4;  DB 84;  Length 386;
  Best Local Similarity   98.6%;  Pred. No. 1.9e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 25

Qy     107 GGGGGGGCC 115
           |||||||||
Db      24 GGGGGGGCC 16


RESULT  13
AQ014512/c
LOCUS       AQ014512      462 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
            genomic survey sequence.
ACCESSION   AQ014512
VERSION     AQ014512.1  GI:3184839
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154425"
                     /db_xref="taxon:9606"
                     /clone="2300D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       89 a    127 c    131 g    115 t
ORIGIN


  Query Match             45.5%;  Score 67.4;  DB 96;  Length 462;
  Best Local Similarity   98.6%;  Pred. No. 1.9e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 42

Qy     107 GGGGGGGCC 115
           |||||||||
Db      41 GGGGGGGCC 33


RESULT  14
B65534/c
LOCUS       B65534        519 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
            genomic survey sequence.
ACCESSION   B65534
VERSION     B65534.1  GI:2639512
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria: Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /db_xref="GDB:7045730"
                     /db_xref="taxon:9606"
                     /clone="2022D24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      110 a    137 c    148 g    124 t
ORIGIN


  Query Match             45.5%;  Score 67.4;  DB 84;  Length 519;
  Best Local Similarity   98.6%;  Pred. No. 2e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 25

Qy     107 GGGGGGGCC 115
           |||||||||
Db      24 GGGGGGGCC 16


RESULT  15
AQ012629
LOCUS       AQ012629      547 bp    DNA             GSS       06-JUN-1998
```

```
DEFINITION  CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
            genomic survey sequence.
ACCESSION   AQ012629
VERSION     AQ012629.1  GI:3185194
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154069"
                     /db_xref="taxon:9606"
                     /clone="2299F1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      154 a    139 c    125 g    129 t
ORIGIN


  Query Match             45.5%;  Score 67.4;  DB 96;  Length 547;
  Best Local Similarity   98.6%;  Pred. No. 2e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      47 TGGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 TAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 132

Qy     107 GGGGGGGCC 115
           |||||||||
Db     133 GGGGGGGCC 141


Search completed: May 29, 2000, 21:13:16
Job time: 36515 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:34 ; Search time 1214.87 Seconds
(without alignments)
934.219 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	73.6	49.4	5639	14	CVU14122	U14122 Cloning vec
c	2	71.2	47.8	259	5	A44281	A44281 Sequence 9
	3	70	47.0	2958	14	ARBLKSM	X52326 pBluescript
	4	70	47.0	2958	14	ARBLKSP	X52331 pBluescript
	5	70	47.0	2961	14	ARBL2KSM	X52329 pBluescript
	6	70	47.0	2961	14	ARBL2KSP	X52327 pBluescript
	7	70	47.0	2964	14	SYNBLKSMV	L08784 BlueScribe
	8	70	47.0	2964	14	SYNBLKSPV	L08785 BlueScribe
c	9	70	47.0	3306	14	SYNPBEN66	D85525 Cloning vec
	10	70	47.0	4144	14	XXU35131	U35131 Plasmid pBS
	11	70	47.0	4267	14	PRS304	U03436 Yeast integ
	12	70	47.0	4289	14	XXU35136	U35136 Plasmid pBS
	13	70	47.0	4373	14	PRS306	U03438 Yeast integ
	14	70	47.0	4443	14	PRS303	U03435 Yeast integ
c	15	70	47.0	4670	14	ASAJ5326	AJ005326 pGAII(+)
	16	70	47.0	4670	14	ASAJ5329	AJ005329 pGAII(-)
	17	70	47.0	4707	14	XXU02374	U02374 Cloning vec
	18	70	47.0	4768	14	XXU25061	U25061 Cloning vec
	19	70	47.0	4783	14	PRS314	U03440 Yeast centr
	20	70	47.0	4887	14	PRS316	U03442 Yeast centr
	21	70	47.0	4950	14	XXU25060	U25060 Cloning vec
	22	70	47.0	4967	14	PRS313	U03439 Yeast centr
	23	70	47.0	5144	14	CVU23751	U23751 Cloning vec
	24	70	47.0	5187	14	U34887	U34887 Yeast integ
	25	70	47.0	5228	14	XXU25059	U25059 Cloning vec
	26	70	47.0	5504	14	PRS305	U03437 Yeast integ
	27	70	47.0	5634	14	CVU14125	U14125 Cloning vec
	28	70	47.0	6018	14	PRS315	U03441 Yeast centr
c	29	70	47.0	6340	14	ASAJ5323	AJ005323 pCP1(-) K
	30	70	47.0	9655	14	SYNPGR8V	AD001531 Cloning v
c	31	68.8	46.2	3240	8	AF015771	AF015771 Magnaport
	32	68	45.6	144	14	SYNPLKRA	M22847 Cloning vec
c	33	68	45.6	147	14	SYNPLKRB	M22848 Cloning vec
c	34	68	45.6	2958	14	ARBLSKM	X52324 pBluescript
c	35	68	45.6	2958	14	ARBLSKP	X52325 pBluescript
c	36	68	45.6	2961	14	ARBL2SKM	X52330 pBluescript
c	37	68	45.6	2961	14	ARBL2SKP	X52328 pBluescript
c	38	68	45.6	2964	14	SYNBLDKPV	L08787 BlueScribe
c	39	68	45.6	2964	14	SYNBLSKMV	L08786 BlueScribe
c	40	68	45.6	3062	14	CVU47947	U47947 Cloning vec
c	41	68	45.6	3228	14	CVE7829	AJ007829 Cloning V
c	42	68	45.6	3345	14	XXU35235	U35235 Plasmid pBS
	43	68	45.6	3357	14	XXU35132	U35132 Plasmid pBS
c	44	68	45.6	3357	14	XXU35132	U35132 Plasmid pBS
	45	68	45.6	3574	14	AF139061	AF139061 Binary ve

ALIGNMENTS

```
RESULT   1
CVU14122/c
LOCUS       CVU14122     5639 bp    DNA   circular  SYN       20-MAY-1995
DEFINITION  Cloning vector pSG935, HIS4-based plasmid, complete sequence.
ACCESSION   U14122
VERSION     U14122.1  GI:537471
KEYWORDS    .
SOURCE      Cloning vector pSG935.
  ORGANISM  Cloning vector pSG935
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 5639)
  AUTHORS   Crane,D.I. and Gould,S.J.
  TITLE     The Pichia pastoris HIS4 gene: nucleotide sequence, creation of a
            non-reverting his4 deletion mutant, and development of HIS4-based
            replicating and integrating plasmids
  JOURNAL   Curr. Genet. 26 (5-6), 443-450 (1994)
  MEDLINE   95179799
REFERENCE   2  (bases 1 to 5639)
  AUTHORS   Gould,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-1994) Stephen J. Gould, Neurogenetics, The
            Kennedy Krieger Institute, 707 North Broadway, Baltimore, MD 21030,
            USA
FEATURES             Location/Qualifiers
     source          1. .5639
                     /organism="Cloning vector pSG935"
                     /plasmid="pSG935"
                     /db_xref="taxon:36585"
                     /lab_host="Pichia pastoris"
     gene            833. .3361
                     /gene="HIS4"
     CDS             833. .3361
                     /gene="HIS4"
                     /codon_start=1
                     /transl_table=11
                     /product="histidinol dehydrogenase"
                     /protein_id="AAA66991.1"
                     /db_xref="GI:537473"
                     /translation="MTFPLLPAYASVAEFDNSLSLVGKAVFPYAADQLHNLIKFTQST
                     ELQVNVQVESSVTEDQFEELIDNLLKLYNNGINEVILDLDLAERVVQRIPGARVIYRT
                     LVDKVASLPANASIAVPFSSPLGDLKSFTNGGSRTVYAFSETAKLVDVTSTVASGIIP
                     IIDARQLTTEYELSEDVKKFPVSEILLASLTTDRPDGLFTTLVADSSNYSLGLVYSSK
                     KSIPEAIRTQTGVYQSRRHGLWYKGATSGATQKLLGIELDCDGDCLKFVVEQTGVGFC
                     HLERTSCFGQSKGLRAMEATLWDRKSNAPEGSYTKRLFDDEVLLNAKIREEADELAEA
                     KSKEDIAWECADLFYFALVRCAKYGVTLDEVERNLDMKSLKVTRRKGDAKPGYTKEQP
                     KEESKPKEVPSEGRIELCKIDVSKASSQEIEDALRRPIQKTEQIMELVKPIVDNVRQN
                     GDKALLELTAKFDGVALKTPVLEAPFPEELMQLPDNVKRAIDLSIDNVRKFHEAQLAE
                     TLQVETCPGVVCSRFARPIEKVGLYIPGGTAILPSTSLMLGVPAKVAGCKEIVFASPP
                     KKDGTLTPEVIYVAHKVGAKCIVLAGGAQAVAAMAYGTETVPKCDKIFGPGNQFVTAA
                     KMMVQNDTSALCSIDMPAGPSEVLVIADKYADPDFVVSDLLSQAEHGIDSQVILLAVD
                     MTDKELARIEDAVHNQAVQLPRVEIVRKCIAHSTTLSVATYEQALEMSNQYAPEHLIL
                     QIENASYVDQVQHAGSVFVGAYSPESCGDYSSGTNHTLPTYGYARQYSGVNTATFQKF
                     ITSQDVTPEGLKHIGQAVMDLAAVEGLDAHRNAVKVRMEKLGLI"
     CDS             complement(4651. .5511)
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="AAA66992.1"
                     /db_xref="GI:537472"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     1473 a    1310 c    1372 g    1484 t
ORIGIN
```

Query Match 49.4%; Score 73.6; DB 14; Length 5639;
Best Local Similarity 78.6%; Pred. No. 6.2e-13;
Matches 88; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCCGGGC 64
           | |||||   || |   || | | |   |   | | | | ||| ||  ||||||||||||
Db     802 AGAAATTCTGAAGCCGTATTTTTATTATCAGTGAGTCAGTCATCAGGAGATCCCCCGGGC 743

Qy      65 TGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 TGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 691
```

```
RESULT   2
A44281/c
LOCUS       A44281         259 bp    DNA             PAT       07-MAR-1997
DEFINITION  Sequence 9 from Patent WO9511300.
ACCESSION   A44281
VERSION     A44281.1  GI:2299115
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Chandley,A.C., Kun,M., Sharkey,A.M., Hargreave,T.B. and Cooke,H.J.
  TITLE     AZOOSPERMIA IDENTIFICATION AND TREATMENT
  JOURNAL   Patent: WO 9511300-A 9 27-APR-1995;
            MEDICAL RES COUNCIL (GB)
COMMENT     Other publication AU 7947794 950508.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /clone_lib="MOUSE GENOMIC LIBRARY"
                     /clone="M3.2"
BASE COUNT       67 a     58 c     63 g     69 t      2 others
ORIGIN
```

Query Match 47.8%; Score 71.2; DB 5; Length 259;
Best Local Similarity 85.9%; Pred. No. 2.6e-12;
Matches 79; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy      43 GCCATGAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAC 102
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GCCTCTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAC 60

Qy     103 CTCGAGGGGGGGCCTAACTAACTAATTTTGTT 134
           |||||||||||||||        |   |||||||
Db      59 CTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 28
```

```
RESULT   3
ARBLKSM
LOCUS       ARBLKSM      2958 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52326
VERSION     X52326.1  GI:58064
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
```

```
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(-)"
BASE COUNT      708 a    754 c    731 g    765 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2958;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy    109 GGGGGGCCTAACTAACTAATTTTGTT 134
          ||||||||       |   |||||||
Db    746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   4
ARBLKSP
LOCUS       ARBLKSP      2958 bp    DNA   circular  SYN       11-JUN-1998
DEFINITION  pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52331
VERSION     X52331.1  GI:58065
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
REFERENCE   4  (bases 1 to 2958)
  AUTHORS   Lampe,D.J., Grant,T.E. and Robertson,H.M.
  TITLE     Factors affecting transposition of the Himar1 mariner transposon in
            vitro
  JOURNAL   Genetics 149 (1), 179-187 (1998)
  MEDLINE   98250682
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(+)"
BASE COUNT      749 a    734 c    751 g    724 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2958;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy    109 GGGGGGCCTAACTAACTAATTTTGTT 134
          ||||||||       |   |||||||
Db    746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   5
ARBL2KSM
LOCUS       ARBL2KSM     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(-) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52329
VERSION     X52329.1  GI:58060
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(-)"
BASE COUNT      706 a    758 c    735 g    762 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2961;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy    109 GGGGGGCCTAACTAACTAATTTTGTT 134
          ||||||||       |   |||||||
Db    746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   6
ARBL2KSP
LOCUS       ARBL2KSP     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(+) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52327
VERSION     X52327.1  GI:58061
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
```

```
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(+)"
BASE COUNT      747 a    738 c    755 g    721 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2961;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   7
SYNBLKSMV
LOCUS       SYNBLKSMV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Minus cloning vector.
ACCESSION   L08784
VERSION     L08784.1  GI:310728
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Minus - Cloning vector
            ENTRY BLUEKSM                #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Minus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0077
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the minus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
               #parent
                 VecBase(3):BlueM13m
               #parent
                 VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                 VecSource(3):bGalKS, GenBank(50):PF1
               #brother
                 VecBase(3):BlueKSp, VecBase(3):BlueSKm
            PARENT
               Features of BlueKSm (2964 bp)
                 residue     source
                  3- 458  5488-5943     phage f1
                 460- 624   236- 400     pUC19c
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-2964   448-2617     pUC19c
               Conflict (cfl) and Mutations (mut): none
            PARENT
               Features of BlueKSm (2964 bp)
                 residue     source
                  3- 458  5488-5943     phage f1
                 460- 624   449- 285 (c) pUC19
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-1031   237-   1 (c) pUC19
                1032-2964  2686- 754 (c) pUC19
               Conflict (cfl) and Mutations (mut): none
            FEATURE
                     643           start of T7-RNA synthesis
                     774        (c) start of T3-RNA synthesis
               1976-2764 789-1 (c) Ap-R; b-lactamase
            POLYLINKER
                 KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
                 SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
               #resistance  Ap
               #indicator beta-galactosidase
            SUMMARY  BlueKSm    #length 2964    #checksum 1589.
FEATURES             Location/Qualifiers
     source          1. .2964
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      708 a    756 c    735 g    765 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2964;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   8
SYNBLKSPV
LOCUS       SYNBLKSPV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Plus cloning vector.
ACCESSION   L08785
```

```
VERSION     L08785.1  GI:310729
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Plus - Cloning vector
            ENTRY BLUEKSP                #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Plus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0078
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the plus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
               #parent
                 VecBase(3):BlueM13p
               #parent
                 VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                 VecSource(3):bGalKS, GenBank(50):PF1
               #brother
                 VecBase(3):BlueKSm, VecBase(3):BlueSKp
            PARENT
               Features of BlueKSp (2964 bp)
                 residue     source
                  3- 458  5943-5488 (c) phage f1
                 460- 624   236- 400     pUC19c
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-2964   448-2617     pUC19c
               Conflict (cfl) and Mutations (mut): none
            PARENT
               Features of BlueKSp (2964 bp)
                 residue     source
                  3- 458  5943-5488 (c) phage f1
                 460- 624   449- 285 (c) pUC19
                 626- 645     1-  20     T7 promoter
                 653- 760   108-   1 (c) BlueKS-polylinker
                 772- 791    20-   1 (c) T3 promoter
                 795-1031   237-   1 (c) pUC19
                1032-2964  2686- 754 (c) pUC19
               Conflict (cfl) and Mutations (mut): none
            FEATURE
                     643          start of T7-RNA synthesis
                     774       (c) start of T3-RNA synthesis
               1976-2764 789-1 (c) Ap-R; b-lactamase
            POLYLINKER
                 KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
                 SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
               #resistance  Ap
               #indicator beta-galactosidase
            SUMMARY  BlueKSp    #length 2964   #checksum 690.
FEATURES                Location/Qualifiers
     source             1. .2964
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
BASE COUNT      750 a    736 c    755 g    723 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 2964;
  Best Local Similarity   88.4%;  Pred. No. 8e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     686 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     746 GGGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   9
SYNPBEN66/c
LOCUS       SYNPBEN66     3306 bp    DNA   circular  SYN       06-FEB-1999
DEFINITION  Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
            beta-lactamase, complete cds.
ACCESSION   D85525
VERSION     D85525.1  GI:1345433
KEYWORDS    plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE      Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
  ORGANISM  Cloning vector pBEN66
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 3306)
  AUTHORS   Yamamotc,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
            Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
            Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
            (Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE   2  (sites)
  AUTHORS   Yamamoto,Y. and Furuyama,J.
  TITLE     One-step disruption by circular DNA in Escherichia coli
  JOURNAL   Unpublished (1996)
FEATURES                Location/Qualifiers
     source             1. .3306
                        /organism="Cloning vector pBEN66"
                        /plasmid="pBEN66"
                        /db_xref="taxon:47800"
                        /lab_host="E.coli"
     promoter           19. .36
                        /note="T3 promoter"
     gene               260. .1075
                        /gene="kan"
     CDS                260. .1075
                        /gene="kan"
                        /note="derived from Tn903; kanamycin resistance gene"
                        /citation=[2]
                        /codon_start=1
                        /transl_table=11
                        /product="aminoglycoside 3'-phosphotransferase"
                        /protein_id="BAA12824.1"
                        /db_xref="GI:1345434"
                        /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                        GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                        GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                        VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
```

```
                RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
    gene        complement(2209. .3069)
                /gene="amp"
    CDS         complement(2209. .3069)
                /gene="amp"
                /note="amipicillin resistance gene"
                /citation=[1]
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase"
                /protein_id="BAA12825.1"
                /db_xref="GI:1345435"
                /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                EIGASLIKHW"
BASE COUNT      854 a    800 c    790 g    862 t
ORIGIN
```

```
  Query Match             47.0%;  Score 70;  DB 14;  Length 3306;
  Best Local Similarity   88.4%;  Pred. No. 8.1e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 63

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db      62 GGGGGGCCCGGTACCCAGCTTTTGTT 37
```

```
RESULT  10
XXU35131
LOCUS       XXU35131     4144 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL159 cloning vector, complete sequence.
ACCESSION   U35131
VERSION     U35131.1  GI:984907
KEYWORDS    .
SOURCE      cloning vector pBSL159.
  ORGANISM  Cloning vector pBSL159
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4144)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4144)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4144
                     /organism="Cloning vector pBSL159"
                     /plasmid="pBSL159"
                     /db_xref="taxon:42704"
     CDS             complement(1129. .1989)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53618.1"
                     /db_xref="GI:984908"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     CDS             complement(join(3704. .4144,1. .219))
                     /codon_start=1
                     /transl_table=11
                     /product="chloramphenicol acetyltransferase"
                     /protein_id="AAC53619.1"
                     /db_xref="GI:984909"
                     /translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAF
                     LKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETF
                     SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNV
                     ANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGG
                     A"
BASE COUNT      1084 a   1009 c    959 g   1092 t
ORIGIN
```

```
  Query Match             47.0%;  Score 70;  DB 14;  Length 4144;
  Best Local Similarity   88.4%;  Pred. No. 8.3e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2725 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2784

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    2785 GGGGGGCCCGGTACCCAGCTTTTGTT 2810
```

```
RESULT  11
PRS304
LOCUS       PRS304       4267 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS304 with TRP1 marker, complete
            sequence.
ACCESSION   U03436
VERSION     U03436.1  GI:416305
KEYWORDS    .
SOURCE      Cloning vector pRS304.
  ORGANISM  Cloning vector pRS304
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4267)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4267)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4267
                     /organism="Cloning vector pRS304"
                     /db_xref="taxon:31827"
BASE COUNT      1128 a    965 c   1092 g   1082 t
ORIGIN
```

```
  Query Match             47.0%;  Score 70;  DB 14;  Length 4267;
  Best Local Similarity   88.4%;  Pred. No. 8.3e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1920 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 1979

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    1980 GGGGGGCCCGGTACCCAGCTTTTGTT 2005
```

```
RESULT  12
XXU35136
LOCUS       XXU35136     4289 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION   U35136
VERSION     U35136.1  GI:984923
KEYWORDS    .
SOURCE      cloning vector pBSL97.
  ORGANISM  Cloning vector pBSL97
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4289)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4289)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4289
                     /organism="Cloning vector pBSL97"
                     /plasmid="pBSL97"
                     /db_xref="taxon:42709"
     CDS             complement(804. .1598)
                     /codon_start=1
                     /transl_table=11
                     /product="neomycin phosphotransferase"
                     /protein_id="AAC53629.1"
                     /db_xref="GI:984924"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
     CDS             complement(3301. .4161)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53630.1"
                     /db_xref="GI:984925"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     1024 a    1157 c    1111 g     997 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 4289;
  Best Local Similarity   88.4%;  Pred. No. 8.3e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2014 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2073

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    2074 GGGGGGCCCGGTACCCAGCTTTTGTT 2099


RESULT  13
PRS306
LOCUS       PRS306       4373 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS306 with URA3 marker, complete
            sequence.
ACCESSION   U03438
VERSION     U03438.1  GI:416307
KEYWORDS    .
SOURCE      Cloning vector pRS306.
  ORGANISM  Cloning vector pRS306
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4373)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4373)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4373
                     /organism="Cloning vector pRS306"
                     /db_xref="taxon:31829"
BASE COUNT     1181 a    987 c    1106 g    1099 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 4373;
  Best Local Similarity   88.4%;  Pred. No. 8.3e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2085

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    2086 GGGGGGCCCGGTACCCAGCTTTTGTT 2111


RESULT  14
PRS303
LOCUS       PRS303       4443 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS303 with HIS3 marker, complete
            sequence.
ACCESSION   U03435
VERSION     U03435.1  GI:416304
KEYWORDS    .
SOURCE      Cloning vector pRS303.
  ORGANISM  Cloning vector pRS303
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4443)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4443)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4443
                     /organism="Cloning vector pRS303"
                     /db_xref="taxon:31826"
BASE COUNT     1149 a    1048 c    1109 g    1137 t
ORIGIN
```

```
  Query Match             47.0%;  Score 70;  DB 14;  Length 4443;
  Best Local Similarity   88.4%;  Pred. No. 8.3e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2106 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2165

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    2166 GGGGGGCCCGGTACCCAGCTTTTGTT 2191


RESULT  15
ASAJ5326/c
LOCUS       ASAJ5326     4670 bp    DNA   circular  SYN       08-FEB-1999
DEFINITION  pGAII(+) KS positive selection cloning vector gltS gene.
ACCESSION   AJ005326
VERSION     AJ005326.1  GI:4028984
KEYWORDS    gltS gene; glutamate permease.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4670)
  AUTHORS   Gal,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
            Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
            6726, HUNGARY
REFERENCE   2  (bases 1 to 4670)
  AUTHORS   Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
            Kalman,M.
  TITLE     A positive selection cloning system based on the gltS gene of
            Escherichia coli
  JOURNAL   Anal. Biochem. 266 (2), 235-238 (1999)
  MEDLINE   99107575
FEATURES             Location/Qualifiers
     source          1. .4670
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     gene            852. .2231
                     /gene="gltS"
     CDS             852. .2231
                     /gene="gltS"
                     /codon_start=1
                     /transl_table=11
                     /product="glutamate permease"
                     /protein_id="CAA06473.1"
                     /db_xref="GI:4028985"
                     /translation="MFHLDTLATLVAATLTLLLGRKLVHSVSFLKKYTIPEPVAGGLL
                     VALALLVLKKSMGWEVNFDMSLRDPLMLAFFATIGLNANIASLRAGGRVVGIFLIVVV
                     GLLVMQNAIGIGMASLLGLDPLMGLLAGSITLSGGHGTGAAWSKLFIERYGFTNATEV
                     AMACATFGLVLGGLIGGPVARQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGDPLV
                     LERPPPRWSSNSPYSESYYARYLVKHSTTPNGIPDDQEVPTAFEKPDVGRMITSLVLI
                     ETIALIAICLTVGKIVAQLLAGTAFELPTFVCVLFVGVILSNGLSIMGFYRVFERAVS
                     VLGNVSLSLFLAMALMGLKLWELASLALPMLAILVVQTIFMALYAIFVTWRMMGKNYD
                     AAVLAAGHCGFGLGATPTAIANMQAITERFGPSHMAFLVVPMVGAFFIDIVNALVIKL
                     YLMLPIFAG"
BASE COUNT     1040 a   1165 c   1231 g   1234 t
ORIGIN


  Query Match             47.0%;  Score 70;  DB 14;  Length 4670;
  Best Local Similarity   88.4%;  Pred. No. 8.4e-12;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 1442

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db    1441 GGGGGGCCCGGTACCCAGCTTTTGTT 1416


Search completed: May 29, 2000, 21:35:38
Job time: 36870 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:27 ; Search time 1446.77 Seconds
(without alignments)
25.767 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	71.2	47.8	259	1	Q87664	Mouse azoospermia
	2	70	47.0	501	1	T04866	Nucleotide analogu
	3	68	45.6	651	1	X20513	Polynucleotide seq
c	4	68	45.6	752	1	V31294	E. coli J96 pathog
c	5	68	45.6	1091	1	T75006	Human endogenous r
c	6	68	45.6	1122	1	T75005	Human endogenous r
c	7	68	45.6	3792	1	Q48463	Plasmid pG+host4 c
	8	68	45.6	4226	1	T39485	Human steroidogene
c	9	68	45.6	5234	1	Q48464	Plasmid pG+host5 c
c	10	68	45.6	6722	1	Q48465	Plasmid pG+host6 c
c	11	68	45.6	10529	1	V09028	Maize fluory2 gene
	12	68	45.6	12814	1	X24730	Swedish-FAD APP ta
	13	68	45.6	15692	1	X24731	London-FAD APP tar
	14	68	45.6	15692	1	X24732	Swedish/London-FAD
	15	68	45.6	15701	1	X24733	Swedish-FAD APP713
c	16	67.8	45.5	545	1	V68808	Human endogenous r
c	17	67.4	45.2	2640	1	Q26664	bDAT. cDNA encodin
c	18	66.2	44.4	3481	1	X02815	DE19731274 Seq ID
c	19	66.2	44.4	3810	1	X02813	DE19731274 Seq ID
c	20	65.4	43.9	5356	1	T43794	Plasmid pRIPHAT (r
	21	65	43.6	84	1	X02800	E. coli biotin DNA
c	22	65	43.6	3465	1	X02814	DE19731274 Seq ID
c	23	65	43.6	3794	1	X02812	DE19731274 Seq ID
	24	65	43.6	6596	1	V57377	Maize female-prefe
c	25	64.6	43.4	545	1	T75010	Human endogenous r
	26	63.4	42.6	685	1	V10190	Stealth virus nucl
	27	63.4	42.6	685	1	V12003	Stealth virus plas
	28	60	40.3	4164	1	T04575	Plasmid pAT-1 sequ
	29	60	40.3	4164	1	V22271	pAT-1 (pSD544). DN
	30	60	40.3	4164	1	V69740	Nucleotide sequenc
	31	60	40.3	5178	1	T49876	pTet-Splice. Nucle
	32	58.8	39.5	2973	1	V64254	Plasmid pKS varian
	33	58.8	39.5	3198	1	T92702	Candida CaRhol gen
	34	58.8	39.5	3198	1	T92869	Candida CaRhol gen
	35	56.4	37.9	459	1	V57250	Clone #4 from muta
	36	56.4	37.9	1949	1	X05602	Nucleotide sequenc
c	37	56	37.6	78	1	X02799	E. coli biotin DNA
c	38	56	37.6	2150	1	T91037	Yeast checkpoint c
c	39	56	37.6	2150	1	X01271	Yeast RAD17 coding
	40	56	37.6	3956	1	V64258	Plasmid pPK13/14 D
	41	56	37.6	4088	1	V64255	Plasmid pPK5/6 DNA
	42	56	37.6	4102	1	V64257	Plasmid pPK9/10 DN
	43	56	37.6	4583	1	V64256	Plasmid pPK7/8 DNA
	44	55.4	37.2	12494	1	V83206	Vector pHP-1 compr
	45	55	36.9	10504	1	V79503	DNA sequence of ex

ALIGNMENTS

RESULT 1
Q87664/c
ID Q87664 standard; DNA; 259 BP.
AC Q87664;
DT 06-MAR-1996 (first entry)
DE Mouse azoospermia factor (AZF) gene partial clone M3.2.
KW Azoospermia factor; AZF; male infertility; YRRM gene;
KW Y-chromosome; RNA recognition motif; ds.
OS Mus sp.
FH Key Location/Qualifiers
FT misc_difference 174
FT /*tag= a
FT /note= "base n at position 174 is not identified
FT in the specification"
FT misc_difference 197
FT /*tag= b
FT /note= "base n at position 197 is not identified
FT in the specification"
PN WO9511300-A2.
PD 27-APR-1995.
PF 24-OCT-1994; G02344.
PR 22-OCT-1993; GB-021857.
PR 07-JUL-1994; GB-013760.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;
DR WPI; 95-170221/22.
PT Nucleic acid encoding the human azoospermia factor, and probes and
PT antibodies specific for the sequence and encoded polypeptide - may
PT be used in the clinical diagnosis of male infertility
PS Disclosure; Fig 5; 40pp; English.
CC A mouse gene partial sequence (Q87664) showing homology to the human
CC azoospermia factor (AZF) YRMM gene was obtd. from mouse genomic phage
CC clone M3.2 isolated from a genomic library screened with MYK (YRRM1)
CC cDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm
CC of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also
CC obtd.
SQ Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;

Query Match 47.8%; Score 71.2; DB 1; Length 259;
Best Local Similarity 85.9%; Pred. No. 7.3e-16;
Matches 79; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy     43 GCCATGAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAC 102
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 GCCTCTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAC 60

Qy    103 CTCGAGGGGGGGCCTAACTAACTAATTTTGTT 134
          ||||||||||||||        |   |||||||
Db     59 CTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 28
```

RESULT 2
T04866
ID T04866 standard; DNA; 501 BP.
AC T04866;

```
DT   28-JAN-1996  (first entry)
DE   Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW   Nucleotide analogue; alkaline phosphatase; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .501
FT                   /*tag= a
FT                   /note= "std. IUPAC codes used"
PN   WO9524505-A.
PD   14-SEP-1995.
PF   07-MAR-1995; G00490.
PR   08-MAR-1994; EP-301636.
PA   (AMSH ) AMERSHAM INT PLC.
PI   Reeve MA,  Robinson PS;
DR   WPI; 95-328290/42.
PT   Modification of residual fluorescence labelled nucleotide analogues  - to
PT   prevent migration in electrophoretic sequencing gel and interference with
PT   base calling of DNA chains
PS   Example; Fig 1; 18pp; English.
CC   The invention concerns the modification of redisual fluorescence
CC   labelled nt analogues to prevent migration in electrophoretic
CC   sequencing gel and interference with base calling of DNA chains. The
CC   modification involves the use of a phosphatase enzyme to remove at
CC   least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC   dye-terminating kt. The template used was 1 microg. of M13 mp8. The
CC   primer was M13 Universal primer. Samples were recovered. Some were
CC   subjected to calf intestinal alkaline phospatase digestion. This
CC   includes T04866. T04867 was a control nt. which was not treated with
CC   alkaline phosphatase. The example shows that alkaline phosphatase
CC   treatment causes removal of the dye-terminator artefacts and allows
CC   for accurate base calling with the ABI analysis software.
SQ   Sequence   501 BP;    98 A;     136 C;    140 G;    121 T;

  Query Match             47.0%;  Score 70;  DB 1;  Length 501;
  Best Local Similarity   88.4%;  Pred. No. 2.4e-15;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 115

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     116 GGGGGGCCCGGTACCCAGCTTTTGTT 141


RESULT   3
X20513
ID   X20513 standard; DNA; 651 BP.
AC   X20513;
DT   05-MAY-1999  (first entry)
DE   Polynucleotide sequence from the genome of Treponema pallidum.
KW   Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW   enzyme production; ds.
OS   Treponema pallidum.
PN   WO9859034-A2.
PD   30-DEC-1998.
PF   23-JUN-1998; U13041.
PR   24-JUN-1997; US-050667.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Fraser CM;
DR   WPI; 99-081273/07.
PT   New isolated Treponema pallidum nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of T. pallidum infections, particularly syphilis
PS   Claim 1; Page 257; 1150pp; English.
CC   X20500-21243 represent polynucleotide sequences from the genome of
CC   Treponema pallidum. The sequences can be used for detection,
CC   diagnosis, characterisation, prevention and therapy for T. pallidum
CC   infections, particularly syphilis. They can also be used for detecting
CC   diseases related to Borrelia infections in animals, and for the
CC   production of biosynthetic products such as enzymes.
SQ   Sequence   651 BP;    180 A;    167 C;    152 G;    149 T;

  Query Match             45.6%;  Score 68;  DB 1;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 136

Qy     109 GGGGGGCC 116
           ||||||||
Db     137 GGGGGGCC 144


RESULT   4
V31294/c
ID   V31294 standard; DNA; 752 BP.
AC   V31294;
DT   01-OCT-1998  (first entry)
DE   E. coli J96 pathogenicity island contig #108.
KW   PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW   PAI V; pheV; vaccine; protective immune response; ds.
OS   Escherichia coli.
PN   WO9822575-A2.
PD   28-MAY-1998.
PF   21-NOV-1997; U21347.
PR   14-OCT-1997; US-061953.
PR   22-NOV-1996; US-031626.
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYWI-) UNIV WISCONSIN.
PI   Choi GH, Dillon PJ, Welch RA;
DR   WPI; 98-312461/27.
PT   New isolated uropathogenic E. coli nucleotide sequences - used to
PT   develop products for the detection of pathogenic E. coli and to
PT   elicit an immune response to pathogenic E. coli
PS   Claim 21; Page 221-222; 250pp; English.
CC   This sequence represents a E. coli strain J96 contig containing
CC   pathogenicity island (PAI) sequences, and represents a nucleic acid
CC   molecule of the invention. PAIs are large fragments of DNA which comprise
CC   pathogenicity determinants. The sequences of the invention are taken from
CC   PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC   on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC   approximately 94 min (at pheR) on the E. coli chromosome and is
CC   approximately 160 kb in size. Antibodies specific to the proteins encoded
CC   by the PAI open reading frames of the invention can be used in kits to
CC   detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC   a protective immune response in an animal to the uropathogenic E. coli
CC   strain J96.
SQ   Sequence   752 BP;    162 A;    213 C;    203 G;    172 T;

  Query Match             45.6%;  Score 68;  DB 1;  Length 752;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 14

Qy     109 GGGGGGCC 116
           ||||||||
Db      13 GGGGGGCC 6


RESULT   5
T75006/c
ID   T75006 standard; DNA; 1091 BP.
AC   T75006;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 6.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
```

```
OS   Human retrovirus.
PN   WO9725431-A1.
PD   17-JUL-1997.
PF   10-JAN-1997; U00398.
PR   10-JAN-1996; US-587329.
PA   (CORI-) CORIXA CORP.
PI   Frudakis TN,  Smith JM;
DR   WPI; 97-384982/35.
PT   Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT   for the prognosis, diagnosis and monitoring of human cancers,
PT   especially breast cancer
PS   Claim 10; Page 31-32; 74pp; English.
CC   Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC   and 11-22 (T75003-10) were obtd. by screening human genomic
CC   libraries using human breast tumour-associated retroviral element
CC   B18Ag1 (see also T75002) as probe.  These non-contiguous sequences
CC   lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC   retrovirus genome (see also T75001).  B18Ag1 and the other
CC   retroviral sequences can be used in genetic vaccines and for the
CC   prognosis, diagnosis and monitoring of human breast cancer.
SQ   Sequence   1091 BP;    79 A;    350 C;    97 G;    248 T;


  Query Match            45.6%;  Score 68;  DB 1;  Length 1091;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    138 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 79

Qy    109 GGGGGGCC 116
           ||||||||
Db     78 GGGGGGCC 71


RESULT   6
T75005/c
ID   T75005 standard; DNA; 1122 BP.
AC   T75005;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 3.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS   Human retrovirus.
PN   WO9725431-A1.
PD   17-JUL-1997.
PF   10-JAN-1997; U00398.
PR   10-JAN-1996; US-587329.
PA   (CORI-) CORIXA CORP.
PI   Frudakis TN,  Smith JM;
DR   WPI; 97-384982/35.
PT   Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT   for the prognosis, diagnosis and monitoring of human cancers,
PT   especially breast cancer
PS   Claim 10; Page 30-31; 74pp; English.
CC   Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC   and 11-22 (T75003-10) were obtd. by screening human genomic
CC   libraries using human breast tumour-associated retroviral element
CC   B18Ag1 (see also T75002) as probe.  These non-contiguous sequences
CC   lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC   retrovirus genome (see also T75001).  B18Ag1 and the other
CC   retroviral sequences can be used in genetic vaccines and for the
CC   prognosis, diagnosis and monitoring of human breast cancer.
SQ   Sequence   1122 BP;    260 A;    316 C;    185 G;    279 T;


  Query Match            45.6%;  Score 68;  DB 1;  Length 1122;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 50

Qy    109 GGGGGGCC 116
           ||||||||
Db     49 GGGGGGCC 42


RESULT   7
Q48463/c
ID   Q48463 standard; DNA; 3792 BP.
AC   Q48463;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host4 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 9; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463). It
CC   is thermosensitive in all hosts tested, including E.coli, and must be
CC   maintained at 28 deg.C.
SQ   Sequence   3792 BP;    1249 A;    552 C;    742 G;    1249 T;


  Query Match            45.6%;  Score 68;  DB 1;  Length 3792;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3542 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483

Qy    109 GGGGGGCC 116
           ||||||||
Db   3482 GGGGGGCC 3475


RESULT   8
T39485
ID   T39485 standard; DNA; 4226 BP.
AC   T39485;
DT   21-MAY-1997  (first entry)
DE   Human steroidogenesis acute regulatory protein genomic DNA.
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; ds.
OS   Homo sapiens.
PN   WO9629338-A1.
PD   26-SEP-1996.
PF   22-MAR-1996; U03896.
PR   23-MAR-1995; US-410540.
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
PI   Lin D,  Miller WL,  Strauss JF;
DR   WPI; 96-443130/44.
PT   Isolated human steroidogenesis acute regulatory protein gene - used
```

```
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
PS   Claim 1; Pages 23-25; 89pp; English.
CC   The present sequence encodes the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc abd pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
SQ   Sequence   4226 BP;    940 A;    1132 C;    1144 G;    984 T;

  Query Match             45.6%;  Score 68;  DB 1;  Length 4226;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4148 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 4207

Qy    109 GGGGGGCC 116
          ||||||||
Db   4208 GGGGGGCC 4215


RESULT   9
Q48464/c
ID   Q48464 standard; DNA; 5234 BP.
AC   Q48464;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host5 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    36. .1496
FT                   /*tag=  a
FT                   /standard_name= ORI
FT                   /note= "origin of replication from pBR322"
FT   misc_feature    2640. .4383
FT                   /*tag=  b
FT                   /note= "from pGK12 (derived from pWV01)"
FT   misc_feature    4384. .4786
FT                   /*tag=  c
FT                   /note= "from pUB110"
FT   misc_feature    4787. .5234
FT                   /*tag=  d
FT                   /note= "from pSK"
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 10; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the 1.4kb AvaI-AlwNI fragment of
CC   pBR322 (containing the origin of replication) was inserted into
CC   NsiI-cleaved pG+host4 to give pG+host5 (Q48464).The pBR322 ORI
CC   allows maintenance of the plasmid in E.coli at 37 deg.C. while the
CC   heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
CC   bacteria.
SQ   Sequence   5234 BP;    1569 A;    973 C;    1133 G;    1559 T;


  Query Match             45.6%;  Score 68;  DB 1;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4984 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 4925

Qy    109 GGGGGGCC 116
          ||||||||
Db   4924 GGGGGGCC 4917


RESULT  10
Q48465/c
ID   Q48465 standard; DNA; 6722 BP.
AC   Q48465;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host6 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Disclosure; Fig 11; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the AvaI-EcoRI fragment of pBR322
CC   (containing the origin of replication and the ampicillin resistance
CC   gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
CC   pG+host6 (Q48466).The pBR322 ORI allows maintenance of the plasmid in
CC   E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC   at 28 deg.C in gram-positive bacteria.
SQ   Sequence   6722 BP;    1961 A;    1278 C;    1503 G;    1980 T;


  Query Match             45.6%;  Score 68;  DB 1;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6472 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 6413

Qy    109 GGGGGGCC 116
          ||||||||
Db   6412 GGGGGGCC 6405
```

```
RESULT  11
V09028/c
ID   V09028 standard; DNA; 10529 BP.
AC   V09028;
DT   20-JUL-1998  (first entry)
DE   Maize fluory2 gene (fl2).
KW   Floury2; fl2 gene; alpha-zein; signal peptide; transgenic plant;
KW   seed; essential amino acid; animal feed; maize; rice; wheat;
KW   barley; millet; sorghum; ds.
OS   Zea mays.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .760
FT                   /*tag= a
FT                   /note= "vector-derived sequence"
FT   promoter        761. .3824
FT                   /*tag= b
FT                   /note= "Claim 14"
FT   CDS             3825. .4613
FT                   /*tag= c
FT   sig_peptide     3825. .3887
FT                   /*tag= d
FT   mat_peptide     3888. .4610
FT                   /*tag= e
PN   WO9802563-A1.
PD   22-JAN-1998.
PF   11-JUL-1997; U11723.
PR   17-JUL-1996; US-021833.
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (UYAR-) UNIV ARIZONA STATE.
PI   Beach L, Coleman CE, Larkins BA;
DR   WPI; 98-110609/10.
DR   P-PSDB; W23977.
PT   Cereal plants containing trans-gene expressing fusion that includes
PT   signal peptide of the fl2 maize gene - and protein having high
PT   content of essential amino acids, producing feeds of improved
PT   nutritional value
PS   Claim 14; Fig 1A-H; 37pp; English.
CC   This is the nucleotide sequence of a clone of the fluory2 (fl2)
CC   gene of maize.  It codes for a 24-kDa alpha-zein protein (see
CC   W23977) that includes a 21-amino acid signal peptide (see W23976)
CC   which targets the alpha-zein to the lumen of the rough endoplasmic
CC   reticulum.  A claimed cereal plant contains a transgene comprising
CC   a first polynucleotide that encodes the fl2 signal peptide and a
CC   second polynucleotide that encodes an agronomically high-value
CC   protein.  Also new are seeds produced by the plants, the transgene
CC   itself, and a transgene that also includes the fl2 promoter.  The
CC   second polynucleotide preferably encodes a protein that has a high
CC   content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC   maize, wheat, rice, barley, millet or sorghum will have increased
CC   contents of these essential amino acids in their seeds (all
CC   claimed).
SQ   Sequence   10529 BP;    2845 A;    2484 C;    2262 G;    2938 T;


  Query Match             45.6%;  Score 68;  DB 1;  Length 10529;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    727 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 668

Qy    109 GGGGGGCC 116
          ||||||||
Db    667 GGGGGGCC 660


RESULT  12
X24730
ID   X24730 standard; DNA; 12814 BP.
AC   X24730;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP targetting vector pMTI-2398.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous reccmbination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2398;
KW   Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1932. .2276
FT                   /*tag= a
FT                   /standard_name= Swedish-FAD APP
FT                   /note= "encodes W97997"
FT   mat_peptide     5369. .6160
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97997, W97998.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 95-103; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP targetting
CC   vector pMTI-2398.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5007,
CC   created using vector pMTI-2398, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced.  The targetted Swedish-FAD
CC   m/hAPP gene expressed m/hAPP protein at levels approaching those
CC   observed for mouse APP in brain.
SQ   Sequence   12814 BP;    1865 A;    1850 C;    1962 G;    1939 T;


  Query Match             45.6%;  Score 68;  DB 1;  Length 12814;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9887 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 9946

Qy    109 GGGGGGCC 116
          ||||||||
Db   9947 GGGGGGCC 9954


RESULT  13
X24731
ID   X24731 standard; DNA; 15692 BP.
AC   X24731;
DT   21-JUN-1999  (first entry)
DE   London-FAD APP targetting vector pMTI-2453.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

```
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2453;
KW   London-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /standard_name= "Swedish-FAD"
FT                   /note= "encodes W97999"
FT   mutation        replace(4990,"")
FT                   /*tag= b
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= c
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W97999.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 104-113; 209pp; English.
CC   This is the nucleotide sequence of London-FAD APP targetting
CC   vector pMTI-2453. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5401,
CC   created using vector pMTI-2453, the London mutation. i.e.
CC   V(717)L, was also introduced.
SQ   Sequence   15692 BP;    2423 A;    2730 C;    2786 G;    2541 T;

  Query Match             45.6%;  Score 68;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12765 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12824

Qy     109 GGGGGGCC 116
           ||||||||
Db   12825 GGGGGGCC 12832


RESULT  14
X24732
ID   X24732 standard; DNA; 15692 BP.
AC   X24732;
DT   21-JUN-1999  (first entry)
DE   Swedish/London-FAD APP targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2454;
KW   London-FAD APP; Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /note= "encodes W98000"
FT   mutation        replace(4849,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4989,"")
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98000.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 114-123; 209pp; English.
CC   This is the nucleotide sequence of Swedish/London-FAD APP targetting
CC   vector pMTI-2454. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5103,
CC   created using vector pMTI-2454, the London mutation. i.e.
CC   V(717)L, and the Swedish mutation, i.e. KM(670,671)NL, were
CC   introduced. The targetted Swedish/London-FAD m/hAPP gene expressed
CC   m/hAPP protein at levels approaching those observed for mouse APP
CC   in brain.
SQ   Sequence   15692 BP;    2423 A;    2731 C;    2783 G;    2542 T;

  Query Match             45.6%;  Score 68;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12765 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12824

Qy     109 GGGGGGCC 116
           ||||||||
Db   12825 GGGGGGCC 12832


RESULT  15
X24733
ID   X24733 standard; DNA; 15701 BP.
AC   X24733;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP713 targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

```
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2455;
KW   Swedish-FAD APP713; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .4983
FT                   /*tag= a
FT                   /note= "encodes W98001"
FT   mutation        replace(4835,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4981,"")
FT                   /*tag= c
FT                   /standard_name= APP713stop
FT   mat_peptide     8232. .9032
FT                   /*tag= d
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98001.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 124-133; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP713 targetting
CC   vector pMTI-2455. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5215,
CC   created using vector pMTI-2455, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced, and also a stop codon (T to
CC   stop at position 714).
SQ   Sequence   15701 BP;    2470 A;    2675 C;    2845 G;    2497 T;


  Query Match             45.6%;  Score 68;  DB 1;  Length 15701;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-14;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12774 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12833

Qy    109 GGGGGGCC 116
          ||||||||
Db  12834 GGGGGGCC 12841


Search completed: May 29, 2000, 21:58:33
Job time: 38181 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:05 ; Search time 621.83 Seconds
(without alignments)
31.146 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	68	45.6	3792	4	US-08-992-334-1	Sequence 1, Appli
c	2	68	45.6	3792	5	US-08-302-752-1	Sequence 1, Appli
	3	68	45.6	4016	2	US-08-410-540-3	Sequence 3, Appli
c	4	68	45.6	5234	4	US-08-992-334-2	Sequence 2, Appli
c	5	68	45.6	5234	5	US-08-302-752-2	Sequence 2, Appli
c	6	68	45.6	6722	4	US-08-992-334-3	Sequence 3, Appli
c	7	68	45.6	6722	5	US-08-302-752-3	Sequence 3, Appli
c	8	67.8	45.5	793	3	US-08-967-101-121	Sequence 121, App
c	9	67.8	45.5	793	4	US-08-592-541-121	Sequence 121, App
	10	65.4	43.9	9318	3	US-08-793-610-6	Sequence 6, Appli
	11	63.4	42.6	685	2	US-08-463-115-56	Sequence 56, Appl
	12	63.4	42.6	685	2	US-08-465-388-56	Sequence 56, Appl
	13	60	40.3	4164	1	US-08-204-675-1	Sequence 1, Appli
	14	60	40.3	4164	3	US-08-660-754-1	Sequence 1, Appli
	15	60	40.3	4164	4	US-08-796-364-1	Sequence 1, Appli
	16	60	40.3	4164	6	PCT-US95-02520-1	Sequence 1, Appli
	17	60	40.3	5178	3	US-08-474-169-2	Sequence 2, Appli
	18	56	37.6	88	1	US-08-144-602B-15	Sequence 15, Appl
c	19	56	37.6	1023	1	US-08-198-446B-1	Sequence 1, Appli
c	20	56	37.6	1023	3	US-08-870-693-1	Sequence 1, Appli
c	21	56	37.6	2150	1	US-08-198-446B-10	Sequence 10, Appl
c	22	56	37.6	2150	3	US-08-870-693-10	Sequence 10, Appl
c	23	55	36.9	834	3	US-08-967-101-113	Sequence 113, App
c	24	55	36.9	834	4	US-08-592-541-113	Sequence 113, App
	25	51.6	34.6	3341	3	US-08-868-577-18	Sequence 18, Appl
c	26	51.2	34.4	1200	1	US-08-011-398B-3	Sequence 3, Appli
c	27	51.2	34.4	1200	2	US-08-464-051-3	Sequence 3, Appli
c	28	51.2	34.4	1200	3	US-08-462-498-3	Sequence 3, Appli
c	29	50	33.6	780	3	US-08-967-101-116	Sequence 116, App
c	30	50	33.6	780	4	US-08-592-541-116	Sequence 116, App
	31	50	33.6	6505	3	US-08-793-610-5	Sequence 5, Appli
c	32	49.8	33.4	6244	1	US-08-076-726-15	Sequence 15, Appl
c	33	49.8	33.4	6244	1	US-08-260-452-8	Sequence 8, Appli
c	34	49.8	33.4	6244	3	US-08-481-970-8	Sequence 8, Appli
c	35	49.8	33.4	6244	4	US-08-897-719-8	Sequence 8, Appli
c	36	48.2	32.3	2308	1	US-08-325-071-62	Sequence 62, Appl
	37	48	32.2	1612	1	US-08-343-733A-1	Sequence 1, Appli
	38	48	32.2	2764	4	US-08-465-971B-1	Sequence 1, Appli
c	39	47.8	32.1	2185	4	US-08-467-948A-3	Sequence 3, Appli
c	40	47.6	31.9	1249	2	US-08-463-115-35	Sequence 35, Appl
c	41	47.6	31.9	1249	2	US-08-465-388-35	Sequence 35, Appl
c	42	47.4	31.8	1997	4	US-08-667-809B-3	Sequence 3, Appli
c	43	47	31.5	2126	3	US-08-789-354-1	Sequence 1, Appli
	44	47	31.5	2605	3	US-08-680-395-4	Sequence 4, Appli
	45	47	31.5	3041	3	US-08-596-300A-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
```

```
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
;    IMMEDIATE SOURCE:
;      CLONE:  pG+host4
US-08-992-334-1


  Query Match             45.6%;  Score 68;  DB 4;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3542 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483

Qy    109 GGGGGGCC 116
          ||||||||
Db   3482 GGGGGGCC 3475


RESULT   2
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-1


  Query Match             45.6%;  Score 68;  DB 5;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3542 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483

Qy    109 GGGGGGCC 116
          ||||||||
Db   3482 GGGGGGCC 3475


RESULT   3
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F.
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,540
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-238/00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4016 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1098..1283
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1620..1733
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2047..2174
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2267..2425
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2567..2751
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2828..2921
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3031..3765
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1433..1434
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2208..2209
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2781..2782
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2956..2957
```

```
;      OTHER INFORMATION:  /note= "interruption of sequence data"
US-08-410-540-3


  Query Match             45.6%;  Score 68;  DB 2;  Length 4016;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3938 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3997

Qy     109 GGGGGGCC 116
           ||||||||
Db    3998 GGGGGGCC 4005


RESULT   4
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-2


  Query Match             45.6%;  Score 68;  DB 4;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4984 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 4925

Qy     109 GGGGGGCC 116
           ||||||||
Db    4924 GGGGGGCC 4917


RESULT   5
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-2


  Query Match             45.6%;  Score 68;  DB 5;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4984 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 4925

Qy     109 GGGGGGCC 116
           ||||||||
Db    4924 GGGGGGCC 4917


RESULT   6
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-3


  Query Match             45.6%;  Score 68;  DB 4;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6472 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 6413

Qy     109 GGGGGGCC 116
           ||||||||
Db    6412 GGGGGGCC 6405


RESULT   7
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-3


  Query Match             45.6%;  Score 68;  DB 5;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-15;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6472 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 6413

Qy     109 GGGGGGCC 116
           ||||||||
Db    6412 GGGGGGCC 6405


RESULT   8
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-967-101-121


  Query Match             45.5%;  Score 67.8;  DB 3;  Length 793;
  Best Local Similarity   86.2%;  Pred. No. 4.5e-15;
  Matches   75;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      48 GAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 107
           | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 GGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 61
```

```
Qy    108 GGGGGGGCCTAACTAACTAATTTTGTT 134
          |||||||||         |   |||||||
Db     60 GGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   9
US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-592-541-121


  Query Match             45.5%;  Score 67.8;  DB 4;  Length 793;
  Best Local Similarity   86.2%;  Pred. No. 4.5e-15;
  Matches   75;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy     48 GAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 107
          | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 GGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 61

Qy    108 GGGGGGGCCTAACTAACTAATTTTGTT 134
          |||||||||         |   |||||||
Db     60 GGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6


  Query Match             43.9%;  Score 65.4;  DB 3;  Length 9318;
  Best Local Similarity   98.5%;  Pred. No. 7.6e-14;
  Matches   66;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     50 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 109
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5654 GGGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 5713

Qy    110 GGGGGCC 116
          |||||||
Db   5714 GGGGGCC 5720


RESULT  11
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
```

```
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,115
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:            Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5703221ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/301
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-463-115-56


  Query Match             42.6%;  Score 63.4;  DB 2;  Length 685;
  Best Local Similarity   94.1%;  Pred. No. 1.5e-13;
  Matches   64;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||
Db     149 AGTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGAG 208

Qy     109 GGGGGGCC 116
           |||||| |
Db     209 GGGGGGNC 216


RESULT  12
US-08-465-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,388
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:            Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5753488ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/300
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-465-388-56


  Query Match             42.6%;  Score 63.4;  DB 2;  Length 685;
  Best Local Similarity   94.1%;  Pred. No. 1.5e-13;
  Matches   64;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||
Db     149 AGTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGAG 208

Qy     109 GGGGGGCC 116
           |||||| |
Db     209 GGGGGGNC 216


RESULT  13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 5677170
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,675
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-204-675-1


  Query Match             40.3%;  Score 60;  DB 1;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-12;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  14
US-08-660-754-1
; Sequence 1, Application US/08660754
; Patent No. 5843772
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/660,754
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NC
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-660-754-1


  Query Match             40.3%;  Score 60;  DB 3;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-12;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  15
US-08-796-364-1
; Sequence 1, Application US/08796364
; Patent No. 5968785
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/796,364
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-796-364-1


  Query Match             40.3%;  Score 60;  DB 4;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-12;
```

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102
```

Search completed: May 29, 2000, 22:09:10
Job time: 38751 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:31 ; Search time 2276.24 Seconds
(without alignments)
292.062 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	100.0	149	28	US-08-935-377-8	Sequence 8, Appli
2	138	92.6	150	28	US-08-935-377-9	Sequence 9, Appli

	3	137	91.9	148	28	US-08-935-377-7	Sequence 7, Appli
	4	131	87.9	145	28	US-08-935-377-6	Sequence 6, Appli
	5	70	47.0	224	13	US-08-377-383-4	Sequence 4, Appli
	6	70	47.0	224	18	US-08-693-573-4	Sequence 4, Appli
	7	69.4	46.6	858	40	US-09-123-912-91	Sequence 91, Appl
	8	69.4	46.6	858	43	US-09-221-107-91	Sequence 91, Appl
	9	69.4	46.6	858	46	US-09-285-479-91	Sequence 91, Appl
c	10	68	45.6	147	38	US-09-072-433-31	Sequence 31, Appl
c	11	68	45.6	157	63	US-60-029-960-595	Sequence 595, App
	12	68	45.6	245	54	US-09-411-999-19130	Sequence 19130, A
	13	68	45.6	381	49	US-09-332-782-8570	Sequence 8570, Ap
	14	68	45.6	381	57	US-09-515-694-8570	Sequence 8570, Ap
	15	68	45.6	393	54	US-09-411-999-13627	Sequence 13627, A
	16	68	45.6	405	49	US-09-332-782-8523	Sequence 8523, Ap
	17	68	45.6	405	57	US-09-515-694-8523	Sequence 8523, Ap
	18	68	45.6	420	49	US-09-332-782-8595	Sequence 8595, Ap
	19	68	45.6	420	57	US-09-515-694-8595	Sequence 8595, Ap
	20	68	45.6	422	49	US-09-332-782-9309	Sequence 9309, Ap
	21	68	45.6	422	57	US-09-515-694-9309	Sequence 9309, Ap
	22	68	45.6	423	54	US-09-411-999-34199	Sequence 34199, A
	23	68	45.6	433	42	US-09-179-473-2334	Sequence 2334, Ap
	24	68	45.6	433	49	US-09-328-351-2334	Sequence 2334, Ap
	25	68	45.6	442	42	US-09-179-473-2495	Sequence 2495, Ap
	26	68	45.6	442	49	US-09-328-351-2495	Sequence 2495, Ap
	27	68	45.6	448	42	US-09-179-473-2255	Sequence 2255, Ap
	28	68	45.6	448	49	US-09-328-351-2255	Sequence 2255, Ap
	29	68	45.6	454	42	US-09-179-473-2100	Sequence 2100, Ap
	30	68	45.6	454	42	US-09-179-473-2360	Sequence 2360, Ap
	31	68	45.6	454	49	US-09-328-351-2100	Sequence 2100, Ap
	32	68	45.6	454	49	US-09-328-351-2360	Sequence 2360, Ap
c	33	68	45.6	752	32	US-08-976-259-108	Sequence 108, App
c	34	68	45.6	752	64	US-60-031-626-108	Sequence 108, App
c	35	68	45.6	827	82	US-60-126-265-996	Sequence 996, App
c	36	68	45.6	2958	37	US-09-040-961-6	Sequence 6, Appli
c	37	68	45.6	2958	41	US-09-130-115-6	Sequence 6, Appli
c	38	68	45.6	2961	15	US-08-446-935-6	Sequence 6, Appli
c	39	67.8	45.5	545	17	US-08-585-392-10	Sequence 10, Appl
c	40	67.8	45.5	545	17	US-08-585-392A-10	Sequence 10, Appl
c	41	67.8	45.5	545	17	US-08-587-329-10	Sequence 10, Appl
c	42	67.8	45.5	545	19	US-08-700-014-10	Sequence 10, Appl
c	43	67.8	45.5	545	23	US-08-838-762-10	Sequence 10, Appl
c	44	67.8	45.5	545	34	US-08-991-789-10	Sequence 10, Appl
c	45	67.8	45.5	545	103	US-09-429-755-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT   1
US-08-935-377-8
; Sequence 8, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-8


  Query Match           100.0%;  Score 149;  DB 28;  Length 149;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-40;
  Matches  149;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60

Qy     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120

Qy    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149
          |||||||||||||||||||||||||||||
Db    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149


RESULT   2
US-08-935-377-9
; Sequence 9, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-9
```

```
  Query Match            92.6%;  Score 138;  DB 28;  Length 150;
  Best Local Similarity  99.3%;  Pred. No. 2.9e-36;
  Matches  149;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGA-GTGGATCCCC 59
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60

Qy     60 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120

Qy    120 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 149
          |||||||||||||||||||||||||||||
Db    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150


RESULT   3
US-08-935-377-7
; Sequence 7, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  148 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-7


  Query Match            91.9%;  Score 137;  DB 28;  Length 148;
  Best Local Similarity  99.3%;  Pred. No. 6.2e-36;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG-GTGGATCCCCC 59

Qy     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 119

Qy    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149
          ||||||||||||||||||||||||||||
Db    120 TAACTAATTTTGTTTTTGTGGGCCCGGCC 148


RESULT   4
US-08-935-377-6
; Sequence 6, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-6


  Query Match            87.9%;  Score 131;  DB 28;  Length 145;
  Best Local Similarity  97.3%;  Pred. No. 6.3e-34;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCC 60
          |||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCC----GTGGATCCCCC 56

Qy     61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAAC 116

Qy    121 TAACTAATTTTGTTTTTGTGGGCCCGGCC 149
          ||||||||||||||||||||||||||||
Db    117 TAACTAATTTTGTTTTTGTGGGCCCGGCC 145


RESULT   5
US-08-377-383-4
; Sequence 4, Application US/08377383
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru, FUKAMI,Tadashi, HIYOSHI,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of
;    TITLE OF INVENTION:  Nucleic Acid
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman Darby & Cushman
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
```

```
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3918
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskett, 3.5 inch, 720Kb storage
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1 or ASCII editors
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/377,383
;       FILING DATE:  25-JAN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP6971/1994
;       FILING DATE:  26-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Glenn J. Perry
;       REGISTRATION NUMBER:  28458
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)861-3000
;       TELEFAX:  (202)822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  224 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-377-383-4


  Query Match             47.0%;  Score 70;  DB 13;  Length 224;
  Best Local Similarity   88.4%;  Pred. No. 2e-13;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 146

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     147 GGGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT   6
US-08-693-573-4
; Sequence 4, Application US/08693573
;  GENERAL INFORMATION:
;     APPLICANT:  HOSOI,Shigeru,
;     APPLICANT:  FUKAMI,Tadashi,
;     APPLICANT:  KOJIMA,Makiko
;     TITLE OF INVENTION:  Method of Determining Base Sequence of                Nucleic
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CUSHMAN DARBY CUSHMAN
;       ADDRESSEE:  INTELLECTUAL PROPERTY GROUP OF
;       ADDRESSEE:  PILLSBURY MADISON & SUTRO, L.L.P.
;       STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005-3918
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.5 inch, 720Kb storage
;       COMPUTER:  IBM PC/XT/AT compatibles
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 5.1 or ASCII editors
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/693,573
;       FILING DATE:  07-AUG-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/377,383
;       FILING DATE:  25-JAN-1995
;       APPLICATION NUMBER:  JP6971/1994
;       FILING DATE:  26-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Thomas G. Wiseman
;       REGISTRATION NUMBER:  35046
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)861-3000
;       TELEFAX:  (202)822-0944
;       TELEX:  6714627 CUSH
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  224 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-693-573-4


  Query Match             47.0%;  Score 70;  DB 18;  Length 224;
  Best Local Similarity   88.4%;  Pred. No. 2e-13;
  Matches   76;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 146

Qy     109 GGGGGGCCTAACTAACTAATTTTGTT 134
           ||||||||       |   |||||||
Db     147 GGGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT   7
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE:  210121.455C1
;  CURRENT APPLICATION NUMBER: US/09/123,912A
;  CURRENT FILING DATE:  1998-07-27
;  PRIOR APPLICATION NUMBER:  09/040,802
;  PRIOR FILING DATE:  1998-03-18
;  NUMBER OF SEQ ID NOS:  114
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 91
;  LENGTH: 858
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: modified_base
;  LOCATION: (570)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (591)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (655)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (664)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (667)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (683)
;  OTHER INFORMATION: Where n is a, c, g or t
;  NAME/KEY: modified_base
;  LOCATION: (711)
;  OTHER INFORMATION: Where n is a, c, g or t
```

```
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (792)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-123-912-91


  Query Match            46.6%;  Score 69.4;  DB 40;  Length 858;
  Best Local Similarity  76.6%;  Pred. No. 4.5e-13;
  Matches   85;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      6 AAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCCGGGCT 65
          | |||   |  |  ||||  |   | | |   |   |   |  ||||||||||||||||
Db    313 ataatctcaccaggagattcaacgaattccaccacactggactagtggatcccccgggct 372

Qy     66 GCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 gcaggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcc 423


RESULT   8
US-09-221-107-91
; Sequence 91, Application US/09221107
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
;  FILE REFERENCE: 210121.455C2
;  CURRENT APPLICATION NUMBER: US/09/221,107
;  CURRENT FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 161
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (570)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (591)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (655)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (664)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (667)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (683)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (711)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (759)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (760)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (765)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (777)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (787)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (792)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (794)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (801)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (804)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (809)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (817)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (820)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-221-107-91
```

```
  Query Match            46.6%;  Score 69.4;  DB 43;  Length 858;
  Best Local Similarity  76.6%;  Pred. No. 4.5e-13;
  Matches   85;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6 AAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCCGGGCT 65
           | |||   |  |  ||||  |   | | |   |   |   |  ||||||||||||||||
Db     313 ataatctcaccaggagattcaacgaattccaccacactggactagtggatcccccgggct 372

Qy      66 GCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 116
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 gcaggaattcgatatcaagcttatcgataccgtcgacctcgaggggggggcc 423


RESULT   9
US-09-285-479-91
; Sequence 91, Application US/09285479
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: LUNG CANCER
;  FILE REFERENCE: 210121.455C3
;  CURRENT APPLICATION NUMBER: US/09/285,479
;  CURRENT FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 172
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 91
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(858)
;   OTHER INFORMATION: n = A,T,C or G
US-09-285-479-91


  Query Match            46.6%;  Score 69.4;  DB 46;  Length 858;
  Best Local Similarity  76.6%;  Pred. No. 4.5e-13;
  Matches   85;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       6 AAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGAGTGGATCCCCCGGGCT 65
           | |||   |  |  ||||  |   | | |   |   |   |  ||||||||||||||||
Db     313 ataatctcaccaggagattcaacgaattccaccacactggactagtggatcccccgggct 372

Qy      66 GCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC 116
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 gcaggaattcgatatcaagcttatcgataccgtcgacctcgaggggggggcc 423


RESULT  10
US-09-072-433-31/c
; Sequence 31, Application US/09072433
;  GENERAL INFORMATION:
;    APPLICANT:  Coschigano, Peter W.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  BIOREMEDIATION
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,433
;      FILING DATE:  04-MAY-1998
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Peter G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  OHU-03344
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  147 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..147
US-09-072-433-31


  Query Match            45.6%;  Score 68;  DB 38;  Length 147;
  Best Local Similarity  100.0%;  Pred. No. 8.2e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 48

Qy     109 GGGGGGCC 116
           ||||||||
Db      47 GGGGGGCC 40


RESULT  11
US-60-029-960-595/c
; Sequence 595, Application US/60029960
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  1649
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/029,960
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  595:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  157 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
```

US-60-029-960-595

Query Match 45.6%; Score 68; DB 63; Length 157;
Best Local Similarity 100.0%; Pred. No. 8.3e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 34

Qy     109 GGGGGGCC 116
           ||||||||
Db      33 GGGGGGCC 26
```

RESULT 12
US-09-411-999-19130
; Sequence 19130, Application US/09411999
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides isolated from plants and
; TITLE OF INVENTION: methods for their use.
; FILE REFERENCE: 1023U
; CURRENT APPLICATION NUMBER: US/09/411,999
; CURRENT FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 37897
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19130
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-411-999-19130

Query Match 45.6%; Score 68; DB 54; Length 245;
Best Local Similarity 100.0%; Pred. No. 9.4e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 agtggatcccccgggctgcaggaattcgatatcaagcttatcgataccgtcgacctcgag 89

Qy     109 GGGGGGCC 116
           ||||||||
Db      90 gggggggcc 97
```

RESULT 13
US-09-332-782-8570
; Sequence 8570, Application US/09332782
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-754CON1
; CURRENT APPLICATION NUMBER: US/09/332,782
; CURRENT FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: US 09/181,430
; EARLIER FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 21027
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8570
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(381)
; OTHER INFORMATION: n = A,T,C or G
US-09-332-782-8570

Query Match 45.6%; Score 68; DB 49; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 agtggatcccccgggctgcaggaattcgatatcaagcttatcgataccgtcgacctcgag 156

Qy     109 GGGGGGCC 116
           ||||||||
Db     157 gggggggcc 164
```

RESULT 14
US-09-515-694-8570
; Sequence 8570, Application US/09515694
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Blun, Linda
; APPLICANT: Cheung, Patrick
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Kofler, Janette
; APPLICANT: Labat, Ivan
; APPLICANT: Lee, Won-Jae
; APPLICANT: Lomeli, Michelle
; APPLICANT: Nguyen, Hong
; APPLICANT: Nguyen, Linh
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Ojeda, Jesse
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sahourieh, Hannah
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tulpule, Mukul
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 754CIP
; CURRENT APPLICATION NUMBER: US/09/515,694
; CURRENT FILING DATE: 2000-02-29
; EARLIER APPLICATION NUMBER: 09/332,782
; EARLIER FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: 09/181,430
; EARLIER FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 21027
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8570
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(381)
;   OTHER INFORMATION: n = A,T,C or G
US-09-515-694-8570


  Query Match             45.6%;  Score 68;  DB 57;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 agtggatcccccgggctgcaggaattcgatatcaagcttatcgataccgtcgacctcgag 156

Qy     109 GGGGGGCC 116
           ||||||||
Db     157 gggggggcc 164


RESULT  15
US-09-411-999-13627
; Sequence 13627, Application US/09411999
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka
;  APPLICANT: Shenk, Michael Andrew
;  TITLE OF INVENTION: Polynucleotides isolated from plants and
;  TITLE OF INVENTION: methods for their use.
;  FILE REFERENCE: 1023U
;  CURRENT APPLICATION NUMBER: US/09/411,999
;  CURRENT FILING DATE: 1999-10-04
;  NUMBER OF SEQ ID NOS: 37897
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13627
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-411-999-13627


  Query Match             45.6%;  Score 68;  DB 54;  Length 393;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 agtggatcccccgggctgcaggaattcgatatcaagcttatcgataccgtcgacctcgag 61

Qy     109 GGGGGGCC 116
           ||||||||
Db      62 gggggggcc 69


Search completed: May 30, 2000, 09:48:32
Job time: 60182 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:16 ; Search time 2192.43 Seconds
(without alignments)
275.462 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	71	47.7	682	96	AQ074693	AQ074693 CIT-HSP-2
c 2	69	46.3	608	96	AQ009167	AQ009167 CIT-HFP-2
c 3	68.4	45.9	528	84	B69688	B69688 CIT978SK-A-
4	68	45.6	108	84	B54292	B54292 CIT-HSP-201
5	68	45.6	141	44	AU037147	AU037147 AU037147
c 6	68	45.6	141	96	B95279	B95279 CIT-HSP-217
7	68	45.6	144	84	B80294	B80294 CIT-HSP-204
c 8	68	45.6	169	96	AQ013255	AQ013255 CIT-HSP-2
9	68	45.6	175	96	AQ041198	AQ041198 CIT-HSP-2
10	68	45.6	259	20	T48593	T48593 ph6f4_19/1T
11	68	45.6	275	27	AA052885	AA052885 T3551 Blo
c 12	68	45.6	386	84	B74162	B74162 CIT-HSP-202
c 13	68	45.6	462	96	AQ014512	AQ014512 CIT-HSP-2
c 14	68	45.6	519	84	B65534	B65534 CIT-HSP-202
15	68	45.6	547	96	AQ012629	AQ012629 CIT-HSP-2
16	68	45.6	552	96	AQ012216	AQ012216 CIT-HSP-2
17	68	45.6	575	96	AQ012221	AQ012221 CIT-HSP-2
c 18	68	45.6	596	96	AQ009165	AQ009165 CIT-HFP-2
c 19	68	45.6	607	84	B64363	B64363 CIT-HSP-202
20	68	45.6	632	84	B70299	B70299 CIT-HSP-202
c 21	68	45.6	638	96	AQ016737	AQ016737 CIT-HSP-2
22	68	45.6	647	84	B68775	B68775 CIT-HSP-205
23	68	45.6	666	84	B80296	B80296 CIT-HSP-204
24	68	45.6	700	96	AQ009153	AQ009153 CIT-HSP-2
c 25	67	45.0	541	96	AQ074695	AQ074695 CIT-HSP-2
c 26	67	45.0	576	38	AA752172	AA752172 96BS0243
27	66.4	44.6	513	96	AQ008437	AQ008437 CIT-HSP-2
c 28	66.4	44.6	541	96	AQ014519	AQ014519 CIT-HSP-2
c 29	63.8	42.8	134	84	B72088	B72088 CIT978SK-13
c 30	63.2	42.4	233	96	AQ012365	AQ012365 CIT-HSP-2
c 31	63	42.3	294	20	T26395	T26395 AB161A4F In
c 32	61.6	41.3	133	84	B26269	B26269 F10B5TF IGF
33	61.4	41.2	298	82	RPU68746	U68746 Riftia pach
34	60	40.3	100	43	AI239372	AI239372 SWOvAFCAP
c 35	58.8	39.5	630	96	AQ003238	AQ003238 RPCI11-19
c 36	57.8	38.8	145	84	B53695	B53695 CIT-HSP-201
c 37	57.4	38.5	603	38	AA752115	AA752115 96BS0066
c 38	57	38.3	144	84	B65391	B65391 CIT-HSP-202
39	56.8	38.1	97	41	AI058116	AI058116 SWAMCAC15
40	56	37.6	471	96	AQ024018	AQ024018 CpGR0359A
41	56	37.6	499	96	AQ023564	AQ023564 CpGR0482A
42	56	37.6	547	96	AQ023548	AQ023548 CpGR0051A
43	56	37.6	626	96	AQ024077	AQ024077 CpGR0415A
c 44	55.4	37.2	502	96	AQ023674	AQ023674 CpGR0007B
c 45	55.2	37.0	149	96	AQ044029	AQ044029 CIT-HSP-2

ALIGNMENTS

```
RESULT  1
AQ074693/c
LOCUS       AQ074693      682 bp    DNA             GSS       20-AUG-1998
DEFINITION  CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23,
            genomic survey sequence.
ACCESSION   AQ074693
VERSION     AQ074693.1  GI:3436811
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2301L23"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      156 a    187 c    187 g    151 t      1 others
ORIGIN


  Query Match           47.7%;  Score 71;  DB 96;  Length 682;
  Best Local Similarity   93.7%;  Pred. No. 8.1e-14;
  Matches   74;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 20

Qy     109 GGGGGGCCTAACTAACTAA 127
           ||||||||  | ||  | ||
Db      19 GGGGGGCCCGAGTACCCAA 1


RESULT  2
AQ009167/c
LOCUS       AQ009167      608 bp    DNA             GSS       27-JUN-1998
DEFINITION  CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13,
            genomic survey sequence.
ACCESSION   AQ009167
VERSION     AQ009167.1  GI:3128599
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
            CIT-HSP-2281N13.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .608
```

```
                     /organism="Homo sapiens"
                     /db_xref="GDB:7147360"
                     /db_xref="taxon:9606"
                     /clone="2281N13"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      141 a    164 c    166 g    137 t
ORIGIN


  Query Match             46.3%;  Score 69;  DB 96;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-13;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 24

Qy    109 GGGGGGCCT 117
           |||||||||
Db     23 GGGGGGCCT 15


RESULT   3
B69688/c
LOCUS       B69688         528 bp    DNA             GSS       18-JUN-1998
DEFINITION  CIT978SK-A-448E10.TVC CIT978SK Homo sapiens genomic clone A-448E10,
            genomic survey sequence.
ACCESSION   B69688
VERSION     B69688.1  GI:2708912
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /db_xref="GDB:5299607"
                     /db_xref="taxon:9606"
                     /clone="A-448E10"
                     /clone_lib="CIT978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
BASE COUNT      118 a    138 c    147 g    125 t
ORIGIN


  Query Match             45.9%;  Score 68.4;  DB 84;  Length 528;
  Best Local Similarity   98.6%;  Pred. No. 5.9e-13;
  Matches   69;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 17

Qy    109 GGGGGGCCTA 118
           |||||||| |
Db     16 GGGGGGCCCA 7


RESULT   4
B54292
LOCUS       B54292         108 bp    DNA             GSS       20-JUN-1998
DEFINITION  CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
            genomic survey sequence.
ACCESSION   B54292
VERSION     B54292.1  GI:2608626
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7044003"
                     /db_xref="taxon:9606"
                     /clone="2017M1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       20 a     35 c     34 g     19 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 84;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 92

Qy    109 GGGGGGCC 116
           ||||||||
Db     93 GGGGGGCC 100


RESULT   5
AU037147
```

```
LOCUS       AU037147      141 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSB532, mRNA sequence.
ACCESSION   AU037147
VERSION     AU037147.1  GI:3983900
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2153012.
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSB532"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       26 a     40 c     42 g     28 t      5 others
ORIGIN
```

```
  Query Match             45.6%;  Score 68;  DB 44;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 127

Qy     109 GGGGGGCC 116
           ||||||||
Db     128 GGGGGGCC 135
```

```
RESULT   6
B95279/c
LOCUS       B95279        141 bp    DNA             GSS       26-JUN-1998
DEFINITION  CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
            genomic survey sequence.
ACCESSION   B95279
VERSION     B95279.1  GI:2977616
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Homo sapiens"
                     /db_xref="GDB:7103988"
                     /db_xref="taxon:9606"
                     /clone="2172N2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       24 a     41 c     41 g     35 t
ORIGIN
```

```
  Query Match             45.6%;  Score 68;  DB 96;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 15

Qy     109 GGGGGGCC 116
           ||||||||
Db      14 GGGGGGCC 7
```

```
RESULT   7
B80294
LOCUS       B80294        144 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
            genomic survey sequence.
ACCESSION   B80294
VERSION     B80294.1  GI:2867317
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2045D19.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="Homo sapiens"
                     /db_xref="GDB:7054558"
                     /db_xref="taxon:9606"
```

```
                     /clone="2045D19"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       34 a     43 c     41 g     26 t
ORIGIN


  Query Match            45.6%;  Score 68;  DB 84;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 128

Qy     109 GGGGGGCC 116
           ||||||||
Db     129 GGGGGGCC 136


RESULT   8
AQ013255/c
LOCUS       AQ013255      169 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
            genomic survey sequence.
ACCESSION   AQ013255
VERSION     AQ013255.1  GI:3185820
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2299C22.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .169
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154018"
                     /db_xref="taxon:9606"
                     /clone="2299C22"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       32 a     47 c     48 g     42 t
ORIGIN


  Query Match            45.6%;  Score 68;  DB 96;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     102 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 43

Qy     109 GGGGGGCC 116
           ||||||||
Db      42 GGGGGGCC 35


RESULT   9
AQ041198
LOCUS       AQ041198      175 bp    DNA             GSS       14-JUL-1998
DEFINITION  CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
            genomic survey sequence.
ACCESSION   AQ041198
VERSION     AQ041198.1  GI:3310469
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .175
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2335D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       43 a     50 c     47 g     35 t
ORIGIN


  Query Match            45.6%;  Score 68;  DB 96;  Length 175;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 131

Qy     109 GGGGGGCC 116
           ||||||||
Db     132 GGGGGGCC 139


RESULT  10
T48593
LOCUS       T48593        259 bp    mRNA            EST       02-FEB-1995
DEFINITION  ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph6f4_19/1TV, mRNA sequence.
```

```
ACCESSION   T48593
VERSION     T48593.1  GI:642793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L.,
            Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
            Sverdlov,E.D.
  TITLE     Outward Alu-primed hncDNA library
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Sverdlov ED
            Structure and Function of Human Genes
            Shemyakin Institute of Bioorganic Chemistry
            16/10 Miklukho-Maklaya, Moscow, 117871, Russia
            Tel: 70953306529
            Fax: 70953306538
            Email: sverd@humgen.siobc.msk.su.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ph6f4_19/1TV"
                     /clone_lib="Outward Alu-primed hncDNA library"
                     /note="Vector: pGEM-3Z; Site_1: EcoR1; Site_2: BamH1; The
                     library was constructed as described in [Obradovic, D.,
                     Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
                     Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
                     Sverdlov, E.D. (1993) Bioorganicheskaya khimia, 20,
                     919-930]. This protocol is based on nested primer strategy
                     using Alu- specific primers (ALN3 and TC-65) that direct
                     the hncDNA synthesis outward of Alu repeats."
BASE COUNT       62 a     66 c     72 g     59 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 20;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 114

Qy     109 GGGGGGCC 116
           ||||||||
Db     115 GGGGGGCC 122


RESULT  11
AA052885
LOCUS       AA052885      275 bp    mRNA            EST       13-SEP-1996
DEFINITION  T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
            brucei cDNA 5', mRNA sequence.
ACCESSION   AA052885
VERSION     AA052885.1  GI:1543898
KEYWORDS    EST.
SOURCE      Trypanosoma brucei brucei.
  ORGANISM  Trypanosoma brucei brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Osanya,A., Murphy,N.B. and Pelle,R.
  TITLE     Trypanosoma brucei brucei cDNAs
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:637578.
            Contact: Osanya A
            Unit 3
            International Livestock Research Institute
            Box 30709, Nairobi, Kenya
            Tel: 254 2 630 743
            Fax: 254 2 631 499
            Email: a.osanya@cgnet.com
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Trypanosoma brucei brucei"
                     /strain="Clone A4"
                     /db_xref="taxon:5702"
                     /clone_lib="Bloodstream form of serodeme ILTat1.1"
                     /note="cDNAs were generated from poly (A+) enriched mRNA
                     prepared from different developmental stages of T.b.brucei
                     by reverse transcription followed by PCR amplification
                     using mini-exon and oligo(dT) primers.  The cDNA generated
                     were utilized in RADES-PCR coupled with differential
                     hybridisations to identify differentially expressed mRNA
                     transcripts. The products which showed to be
                     differentially expressed were cloned pGEM -T vector.
                     These differentially expressed mRNA transcripts were (are
                     being) sequenced to generate differentially expressed
                     sequence tags."
BASE COUNT       61 a     84 c     74 g     56 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 27;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107

Qy     109 GGGGGGCC 116
           ||||||||
Db     108 GGGGGGCC 115


RESULT  12
B74162/c
LOCUS       B74162        386 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
            genomic survey sequence.
ACCESSION   B74162
VERSION     B74162.1  GI:2769849
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 386)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2028N1.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .386
                     /organism="Homo sapiens"
                     /db_xref="GDB:7048251"
```

```
                     /db_xref="taxon:9606"
                     /clone="2028N1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       81 a     98 c    109 g     98 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 84;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 24

Qy     109 GGGGGGCC 116
           ||||||||
Db      23 GGGGGGCC 16


RESULT  13
AQ014512/c
LOCUS       AQ014512      462 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
            genomic survey sequence.
ACCESSION   AQ014512
VERSION     AQ014512.1  GI:3184839
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154425"
                     /db_xref="taxon:9606"
                     /clone="2300D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       89 a    127 c    131 g    115 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 96;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     100 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 41

Qy     109 GGGGGGCC 116
           ||||||||
Db      40 GGGGGGCC 33


RESULT  14
B65534/c
LOCUS       B65534        519 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
            genomic survey sequence.
ACCESSION   B65534
VERSION     B65534.1  GI:2639512
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /db_xref="GDB:7045730"
                     /db_xref="taxon:9606"
                     /clone="2022D24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      110 a    137 c    148 g    124 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 84;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 8e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 24

Qy     109 GGGGGGCC 116
           ||||||||
Db      23 GGGGGGCC 16


RESULT  15
AQ012629
LOCUS       AQ012629      547 bp    DNA             GSS       06-JUN-1998
```

```
DEFINITION  CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
            genomic survey sequence.
ACCESSION   AQ012629
VERSION     AQ012629.1  GI:3185194
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154069"
                     /db_xref="taxon:9606"
                     /clone="2299F1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      154 a    139 c    125 g    129 t
ORIGIN


  Query Match             45.6%;  Score 68;  DB 96;  Length 547;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      49 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 AGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 133

Qy     109 GGGGGGCC 116
           ||||||||
Db     134 GGGGGGCC 141


Search completed: May 29, 2000, 21:13:16
Job time: 36515 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:38 ; Search time 1214.87 Seconds
(without alignments)
940.489 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	69.8	46.5	793	5	AR060142	AR060142 Sequence
c	2	69	46.0	259	5	A44281	A44281 Sequence 9
	3	69	46.0	2958	14	ARBLKSM	X52326 pBluescript
	4	69	46.0	2958	14	ARBLKSP	X52331 pBluescript
	5	69	46.0	2961	14	ARBL2KSM	X52329 pBluescript
	6	69	46.0	2961	14	ARBL2KSP	X52327 pBluescript
	7	69	46.0	2964	14	SYNBLKSMV	L08784 BlueScribe
	8	69	46.0	2964	14	SYNBLKSPV	L08785 BlueScribe
c	9	69	46.0	3306	14	SYNPBEN66	D85525 Cloning vec
	10	69	46.0	4144	14	XXU35131	U35131 Plasmid pBS
	11	69	46.0	4267	14	PRS304	U03436 Yeast integ
	12	69	46.0	4289	14	XXU35136	U35136 Plasmid pBS
	13	69	46.0	4373	14	PRS306	U03438 Yeast integ
	14	69	46.0	4443	14	PRS303	U03435 Yeast integ
c	15	69	46.0	4670	14	ASAJ5326	AJ005326 pGAII(+)
	16	69	46.0	4670	14	ASAJ5329	AJ005329 pGAII(-)
	17	69	46.0	4707	14	XXU02374	U02374 Cloning vec
	18	69	46.0	4768	14	XXU25061	U25061 Cloning vec
	19	69	46.0	4783	14	PRS314	U03440 Yeast centr
	20	69	46.0	4887	14	PRS316	U03442 Yeast centr
	21	69	46.0	4950	14	XXU25060	U25060 Cloning vec
	22	69	46.0	4967	14	PRS313	U03439 Yeast centr
	23	69	46.0	5144	14	CVU23751	U23751 Cloning vec
	24	69	46.0	5187	14	U34887	U34887 Yeast integ
	25	69	46.0	5228	14	XXU25059	U25059 Cloning vec
	26	69	46.0	5504	14	PRS305	U03437 Yeast integ
	27	69	46.0	5634	14	CVU14125	U14125 Cloning vec
	28	69	46.0	6018	14	PRS315	U03441 Yeast centr
c	29	69	46.0	6340	14	ASAJ5323	AJ005323 pCP1(-) K
	30	69	46.0	9655	14	SYNPGR8V	AD001531 Cloning v
c	31	68	45.3	2998	14	XXU43955	U43955 Expression
c	32	67.6	45.1	400	13	G16589	G16589 human STS S
c	33	67.4	44.9	3240	8	AF015771	AF015771 Magnaport
	34	67	44.7	144	14	SYNPLKRA	M22847 Cloning vec
c	35	67	44.7	147	14	SYNPLKRB	M22848 Cloning vec
c	36	67	44.7	2958	14	ARBLSKM	X52324 pBluescript
c	37	67	44.7	2958	14	ARBLSKP	X52325 pBluescript
c	38	67	44.7	2961	14	ARBL2SKM	X52330 pBluescript
c	39	67	44.7	2961	14	ARBL2SKP	X52328 pBluescript
c	40	67	44.7	2964	14	SYNBLDKPV	L08787 BlueScribe
c	41	67	44.7	2964	14	SYNBLSKMV	L08786 BlueScribe
c	42	67	44.7	3062	14	CVU47947	U47947 Cloning vec
c	43	67	44.7	3228	14	CVE7829	AJ007829 Cloning V
c	44	67	44.7	3345	14	XXU35235	U35235 Plasmid pBS
	45	67	44.7	3357	14	XXU35132	U35132 Plasmid pBS

ALIGNMENTS

```
RESULT   1
AR060142/c
LOCUS       AR060142      793 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 121 from patent US 5840540.
ACCESSION   AR060142
VERSION     AR060142.1  GI:5986592
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   St. George-Hyslop,P.H., Rommens,J.M. and Fraser,P.E.
  TITLE     Nucleic acids encoding presenilin II
  JOURNAL   Patent: US 5840540-A 121 24-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .793
                     /organism="unknown"
BASE COUNT      162 a    242 c    146 g    229 t     14 others
ORIGIN

  Query Match             46.5%;  Score 69.8;  DB 5;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 1.1e-12;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   2
A44281/c
LOCUS       A44281        259 bp    DNA             PAT       07-MAR-1997
DEFINITION  Sequence 9 from Patent WO9511300.
ACCESSION   A44281
VERSION     A44281.1  GI:2299115
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Chandley,A.C., Kun,M., Sharkey,A.M., Hargreave,T.B. and Cooke,H.J.
  TITLE     AZOOSPERMIA IDENTIFICATION AND TREATMENT
  JOURNAL   Patent: WO 9511300-A 9 27-APR-1995;
            MEDICAL RES COUNCIL (GB)
COMMENT     Other publication AU 7947794 950508.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /clone_lib="MOUSE GENOMIC LIBRARY"
                     /clone="M3.2"
BASE COUNT       67 a     58 c     63 g     69 t      2 others
ORIGIN

  Query Match             46.0%;  Score 69;  DB 5;  Length 259;
  Best Local Similarity   88.2%;  Pred. No. 1.7e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 53

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db      52 GGGGGCCCGGTACCCAGCTTTTGTT 28


RESULT   3
ARBLKSM
LOCUS       ARBLKSM       2958 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52326
VERSION     X52326.1  GI:58064
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(-)"
BASE COUNT      708 a    754 c    731 g    765 t
ORIGIN

  Query Match             46.0%;  Score 69;  DB 14;  Length 2958;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   4
ARBLKSP
LOCUS       ARBLKSP       2958 bp    DNA   circular  SYN       11-JUN-1998
DEFINITION  pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION   X52331
VERSION     X52331.1  GI:58065
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2958)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2958)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2958)
```

```
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
REFERENCE   4  (bases 1 to 2958)
  AUTHORS   Lampe,D.J., Grant,T.E. and Robertson,H.M.
  TITLE     Factors affecting transposition of the Himar1 mariner transposon in
            vitro
  JOURNAL   Genetics 149 (1), 179-187 (1998)
  MEDLINE   98250682
FEATURES             Location/Qualifiers
     source          1. .2958
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2958
                     /note="phagemid pBluescript KS(+)"
BASE COUNT      749 a    734 c    751 g    724 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 2958;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   5
ARBL2KSM
LOCUS       ARBL2KSM     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(-) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52329
VERSION     X52329.1  GI:58060
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(-)"
BASE COUNT      706 a    758 c    735 g    762 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 2961;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   6
ARBL2KSP
LOCUS       ARBL2KSP     2961 bp    DNA   circular  SYN       10-MAY-1995
DEFINITION  pBluescript II KS(+) vector DNA, phagemid excised from lambda
            ZAPII.
ACCESSION   X52327
VERSION     X52327.1  GI:58061
KEYWORDS    artificial sequence; cloning vector; expression vector; vector.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2961)
  AUTHORS   Thomas,E.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
            11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE   2  (bases 1 to 2961)
  AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
  TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
            excision properties
  JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
  MEDLINE   88319944
REFERENCE   3  (bases 1 to 2961)
  AUTHORS   Alting-Mees,M.A. and Short,J.M.
  TITLE     pBluescript II: gene mapping vectors
  JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
  MEDLINE   90067967
FEATURES             Location/Qualifiers
     source          1. .2961
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     misc_feature    1. .2961
                     /note="phagemid pBluescriptII KS(+)"
BASE COUNT      747 a    738 c    755 g    721 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 2961;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   7
SYNBLKSMV
LOCUS       SYNBLKSMV    2964 bp    DNA   circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Minus cloning vector.
ACCESSION   L08784
VERSION     L08784.1  GI:310728
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
```

```
JOURNAL    Unpublished (1991)
COMMENT    These data and their annotation were supplied to GenBank by Will
           Gilbert under the auspices of the GenBank Currator Program.
           BlueScribe KS Minus - Cloning vector
           ENTRY BLUEKSM                #TYPE DNA CIRCULAR
           TITLE BlueScribe KS Minus - Cloning vector
           DATE    28-JAN-1987
              #sequence  02-FEB-1987
              #sequence  04-MAR-1987
              #sequence  03-APR-1987
           ACCESSION VB0077
           SOURCE    artificial
           REFERENCE
              #number 1
              #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
           J.
              #journal Gene (1987) in press
              #citation Sequence data from StrataGene
              #comment  sequence correction according to Stratagene COMMENT
              Obtained from StrataGene on floppy disc.
              Revised 02-FEB-1987 by F. Pfeiffer:
              1409/10 'AT' to 'TA' to match revised sequence of PBR322
              Revised 4-MAR-1987 to match sequence of pUC19 on request
              of Stratagene
              Polylinker region revised 03-APR-1987 according to Stratagene
           COMMENT
              The stand shown corresponds to pUC19c.
              As in the published sequence of pUC19c, The M13mp19 lacZ region
              is on the complementary strand.
           COMMENT
              This vector contains the f1 origin so that the minus strand
              can be obtained upon f1 superinfection.
           KEYWORDS
           CROSSREFERENCE
              #parent
                VecBase(3):BlueM13m
              #parent
                VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                VecSource(3):bGalKS, GenBank(50):PF1
              #brother
                VecBase(3):BlueKSp, VecBase(3):BlueSKm
           PARENT
              Features of BlueKSm (2964 bp)
                residue     source
                 3- 458  5488-5943     phage f1
                460- 624   236- 400     pUC19c
                626- 645     1-  20     T7 promoter
                653- 760   108-   1 (c) BlueKS-polylinker
                772- 791    20-   1 (c) T3 promoter
                795-2964   448-2617     pUC19c
              Conflict (cfl) and Mutations (mut): none
           PARENT
              Features of BlueKSm (2964 bp)
                residue     source
                 3- 458  5488-5943     phage f1
                460- 624   449- 285 (c) pUC19
                626- 645     1-  20     T7 promoter
                653- 760   108-   1 (c) BlueKS-polylinker
                772- 791    20-   1 (c) T3 promoter
                795-1031   237-   1 (c) pUC19
               1032-2964  2686- 754 (c) pUC19
              Conflict (cfl) and Mutations (mut): none
           FEATURE
                    643           start of T7-RNA synthesis
                    774        (c) start of T3-RNA synthesis
              1976-2764 789-1 (c) Ap-R; b-lactamase
           POLYLINKER
                KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
                SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
              #resistance  Ap
              #indicator beta-galactosidase
           SUMMARY  BlueKSm    #length 2964    #checksum 1589.
FEATURES             Location/Qualifiers
     source          1. .2964
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      708 a    756 c    735 g    765 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 2964;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   8
SYNBLKSPV
LOCUS       SYNBLKSPV    2964 bp    DNA    circular  SYN       26-JUL-1993
DEFINITION  BlueScribe KS Plus cloning vector.
ACCESSION   L08785
VERSION     L08785.1  GI:310729
KEYWORDS    .
SOURCE      Synthetic construct DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 2964)
  AUTHORS   Gilbert,W.
  TITLE     Obtained from VecBase 3.0
  JOURNAL   Unpublished (1991)
COMMENT     These data and their annotation were supplied to GenBank by Will
            Gilbert under the auspices of the GenBank Currator Program.
            BlueScribe KS Plus - Cloning vector
            ENTRY BLUEKSP                #TYPE DNA CIRCULAR
            TITLE BlueScribe KS Plus - Cloning vector
            DATE    28-JAN-1987
               #sequence  02-FEB-1987
               #sequence  04-MAR-1987
               #sequence  03-APR-1987
            ACCESSION VB0078
            SOURCE    artificial
            REFERENCE
               #number 1
               #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
            J.
               #journal Gene (1987) in press
               #citation Sequence data from StrataGene
               #comment  sequence correction according to Stratagene COMMENT
               Obtained from StrataGene on floppy disc.
               Revised 02-FEB-1987 by F. Pfeiffer:
               1409/10 'AT' to 'TA' to match revised sequence of PBR322
               Revised 4-MAR-1987 to match sequence of pUC19 on request
               of Stratagene
               Polylinker region revised 03-APR-1987 according to Stratagene
            COMMENT
               The stand shown corresponds to pUC19c.
               As in the published sequence of pUC19c, The M13mp19 lacZ region
               is on the complementary strand.
            COMMENT
               This vector contains the f1 origin so that the plus strand
               can be obtained upon f1 superinfection.
            KEYWORDS
            CROSSREFERENCE
               #parent
                 VecBase(3):BlueM13p
               #parent
                 VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
                 VecSource(3):bGalKS, GenBank(50):PF1
               #brother
```

```
          VecBase(3):BlueKSm, VecBase(3):BlueSKp
       PARENT
          Features of BlueKSp (2964 bp)
            residue     source
              3- 458  5943-5488 (c) phage f1
            460- 624   236- 400     pUC19c
            626- 645     1-  20     T7 promoter
            653- 760   108-   1 (c) BlueKS-polylinker
            772- 791    20-   1 (c) T3 promoter
            795-2964   448-2617     pUC19c
          Conflict (cfl) and Mutations (mut): none
       PARENT
          Features of BlueKSp (2964 bp)
            residue     source
              3- 458  5943-5488 (c) phage f1
            460- 624   449- 285 (c) pUC19
            626- 645     1-  20     T7 promoter
            653- 760   108-   1 (c) BlueKS-polylinker
            772- 791    20-   1 (c) T3 promoter
            795-1031   237-   1 (c) pUC19
           1032-2964  2686- 754 (c) pUC19
          Conflict (cfl) and Mutations (mut): none
       FEATURE
                643           start of T7-RNA synthesis
                774       (c) start of T3-RNA synthesis
          1976-2764 789-1 (c) Ap-R; b-lactamase
       POLYLINKER
            KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
            SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
          #resistance  Ap
          #indicator beta-galactosidase
       SUMMARY  BlueKSp    #length 2964   #checksum 690.
FEATURES             Location/Qualifiers
     source          1. .2964
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      750 a    736 c    755 g    723 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 2964;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     747 GGGGGCCCGGTACCCAGCTTTTGTT 771


RESULT   9
SYNPBEN66/c
LOCUS       SYNPBEN66     3306 bp    DNA   circular  SYN       06-FEB-1999
DEFINITION  Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
            beta-lactamase, complete cds.
ACCESSION   D85525
VERSION     D85525.1  GI:1345433
KEYWORDS    plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE      Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
  ORGANISM  Cloning vector pBEN66
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 3306)
  AUTHORS   Yamamoto,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
            Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
            Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
            (Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE   2  (sites)
  AUTHORS   Yamamoto,Y. and Furuyama,J.
  TITLE     One-step disruption by circular DNA in Escherichia coli
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1. .3306
                     /organism="Cloning vector pBEN66"
                     /plasmid="pBEN66"
                     /db_xref="taxon:47800"
                     /lab_host="E.coli"
     promoter        19. .36
                     /note="T3 promoter"
     gene            260. .1075
                     /gene="kan"
     CDS             260. .1075
                     /gene="kan"
                     /note="derived from Tn903; kanamycin resistance gene"
                     /citation=[2]
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside 3'-phosphotransferase"
                     /protein_id="BAA12824.1"
                     /db_xref="GI:1345434"
                     /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                     GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                     GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                     VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
                     RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
     gene            complement(2209. .3069)
                     /gene="amp"
     CDS             complement(2209. .3069)
                     /gene="amp"
                     /note="amipicillin resistance gene"
                     /citation=[1]
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="BAA12825.1"
                     /db_xref="GI:1345435"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT      854 a    800 c    790 g    862 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 3306;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 62

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |  |||||||
Db      61 GGGGGCCCGGTACCCAGCTTTTGTT 37


RESULT  10
XXU35131
LOCUS       XXU35131     4144 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL159 cloning vector, complete sequence.
ACCESSION   U35131
VERSION     U35131.1  GI:984907
KEYWORDS    .
SOURCE      cloning vector pBSL159.
  ORGANISM  Cloning vector pBSL159
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4144)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
```

```
          for Escherichia coli vector construction and in vitro
          deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4144)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
          Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4144
                     /organism="Cloning vector pBSL159"
                     /plasmid="pBSL159"
                     /db_xref="taxon:42704"
     CDS             complement(1129. .1989)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53618.1"
                     /db_xref="GI:984908"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     CDS             complement(join(3704. .4144,1. .219))
                     /codon_start=1
                     /transl_table=11
                     /product="chloramphenicol acetyltransferase"
                     /protein_id="AAC53619.1"
                     /db_xref="GI:984909"
                     /translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAF
                     LKTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETF
                     SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFIENMFFVSANPWVSFTSFDLNV
                     ANMDNFFAPVFTMGKYYTQGDKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGG
                     A"
BASE COUNT     1084 a   1009 c    959 g   1092 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 4144;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2726 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2785

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db    2786 GGGGGCCCGGTACCCAGCTTTTGTT 2810


RESULT  11
PRS304
LOCUS       PRS304       4267 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS304 with TRP1 marker, complete
            sequence.
ACCESSION   U03436
VERSION     U03436.1  GI:416305
KEYWORDS    .
SOURCE      Cloning vector pRS304.
  ORGANISM  Cloning vector pRS304
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4267)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4267)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4267
                     /organism="Cloning vector pRS304"
                     /db_xref="taxon:31827"
BASE COUNT     1128 a    965 c   1092 g   1082 t
ORIGIN


  Query Match             46.0%;  Score 69;  DB 14;  Length 4267;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1980

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db    1981 GGGGGCCCGGTACCCAGCTTTTGTT 2005


RESULT  12
XXU35136
LOCUS       XXU35136     4289 bp    DNA   circular  SYN       26-SEP-1995
DEFINITION  Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION   U35136
VERSION     U35136.1  GI:984923
KEYWORDS    .
SOURCE      cloning vector pBSL97.
  ORGANISM  Cloning vector pBSL97
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4289)
  AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
  TITLE     Improved antibiotic-resistance gene cassettes and omega elements
            for Escherichia coli vector construction and in vitro
            deletion/insertion mutagenesis
  JOURNAL   Gene 160 (1), 63-67 (1995)
  MEDLINE   95354958
REFERENCE   2  (bases 1 to 4289)
  AUTHORS   Hengen,P.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
            Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES             Location/Qualifiers
     source          1. .4289
                     /organism="Cloning vector pBSL97"
                     /plasmid="pBSL97"
                     /db_xref="taxon:42709"
     CDS             complement(804. .1598)
                     /codon_start=1
                     /transl_table=11
                     /product="neomycin phosphotransferase"
                     /protein_id="AAC53629.1"
                     /db_xref="GI:984924"
                     /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
                     PVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVLDVVTEAGRDWLLLGEVPGQDL
                     LSSHLAPAEKVSIMADAMRRLHTLDPATCPFDHQAKHRIERARTRMEAGLVDQDDLDE
                     EHQGLAPAELFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRY
                     QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
     CDS             complement(3301. .4161)
                     /EC_number="3.5.2.6"
                     /codon_start=1
                     /transl_table=11
                     /function="ampicillin resistance"
                     /product="beta-lactamase"
                     /protein_id="AAC53630.1"
                     /db_xref="GI:984925"
```

```
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
BASE COUNT     1024 a   1157 c   1111 g    997 t
ORIGIN


  Query Match           46.0%;  Score 69;  DB 14;  Length 4289;
  Best Local Similarity   88.2%;  Pred. No. 2.3e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2015 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2074

Qy    111 GGGGGCCTAACTAACTAATTTTGTT 135
          |||||||       |   |||||||
Db   2075 GGGGGCCCGGTACCCAGCTTTTGTT 2099


RESULT  13
PRS306
LOCUS       PRS306       4373 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS306 with URA3 marker, complete
            sequence.
ACCESSION   U03438
VERSION     U03438.1  GI:416307
KEYWORDS    .
SOURCE      Cloning vector pRS306.
  ORGANISM  Cloning vector pRS306
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4373)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4373)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4373
                     /organism="Cloning vector pRS306"
                     /db_xref="taxon:31829"
BASE COUNT     1181 a    987 c   1106 g   1099 t
ORIGIN


  Query Match           46.0%;  Score 69;  DB 14;  Length 4373;
  Best Local Similarity   88.2%;  Pred. No. 2.4e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2086

Qy    111 GGGGGCCTAACTAACTAATTTTGTT 135
          |||||||       |   |||||||
Db   2087 GGGGGCCCGGTACCCAGCTTTTGTT 2111


RESULT  14
PRS303
LOCUS       PRS303       4443 bp    DNA   circular  SYN       14-SEP-1995
DEFINITION  Yeast integrative vector pRS303 with HIS3 marker, complete
            sequence.
ACCESSION   U03435
VERSION     U03435.1  GI:416304
KEYWORDS    .
SOURCE      Cloning vector pRS303.
  ORGANISM  Cloning vector pRS303
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4443)
  AUTHORS   Sikorski,R.S. and Hieter,P.
  TITLE     A system of shuttle vectors and yeast host strains designed for
            efficient manipulation of DNA in Saccharomyces cerevisiae
  JOURNAL   Genetics 122 (1), 19-27 (1989)
  MEDLINE   89276910
REFERENCE   2  (bases 1 to 4443)
  AUTHORS   Stillman,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
            and Molecular Biology, University of Utah Medical Center, Salt Lake
            City, UT 84132, USA
FEATURES             Location/Qualifiers
     source          1. .4443
                     /organism="Cloning vector pRS303"
                     /db_xref="taxon:31826"
BASE COUNT     1149 a   1048 c   1109 g   1137 t
ORIGIN


  Query Match           46.0%;  Score 69;  DB 14;  Length 4443;
  Best Local Similarity   88.2%;  Pred. No. 2.4e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2107 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166

Qy    111 GGGGGCCTAACTAACTAATTTTGTT 135
          |||||||       |   |||||||
Db   2167 GGGGGCCCGGTACCCAGCTTTTGTT 2191


RESULT  15
ASAJ5326/c
LOCUS       ASAJ5326     4670 bp    DNA   circular  SYN       08-FEB-1999
DEFINITION  pGAII(+) KS positive selection cloning vector gltS gene.
ACCESSION   AJ005326
VERSION     AJ005326.1  GI:4028984
KEYWORDS    gltS gene; glutamate permease.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4670)
  AUTHORS   Gal,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
            Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
            6726, HUNGARY
REFERENCE   2  (bases 1 to 4670)
  AUTHORS   Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
            Kalman,M.
  TITLE     A positive selection cloning system based on the gltS gene of
            Escherichia coli
  JOURNAL   Anal. Biochem. 266 (2), 235-238 (1999)
  MEDLINE   99107575
FEATURES             Location/Qualifiers
     source          1. .4670
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     gene            852. .2231
                     /gene="gltS"
     CDS             852. .2231
                     /gene="gltS"
                     /codon_start=1
                     /transl_table=11
                     /product="glutamate permease"
                     /protein_id="CAA06473.1"
```

```
                     /db_xref="GI:4028985"
                     /translation="MFHLDTLATLVAATLTLLLGRKLVHSVSFLKKYTIPEPVAGGLL
                     VALALLVLKKSMGWEVNFDMSLRDPLMLAFFATIGLNANIASLRAGGRVVGIFLIVVV
                     GLLVMQNAIGIGMASLLGLDPLMGLLAGSITLSGGHGTGAAWSKLFIERYGFTNATEV
                     AMACATFGLVLGGLIGGPVARQLTLTKGNKSWVPGPPSRSTVSISLISNSCSPGDPLV
                     LERPPPRWSSNSPYSESYYARYLVKHSTTPNGIPDDQEVPTAFEKPDVGRMITSLVLI
                     ETIALIAICLTVGKIVAQLLAGTAFELPTFVCVLFVGVILSNGLSIMGFYRVFERAVS
                     VLGNVSLSLFLAMALMGLKLWELASLALPMLAILVVQTIFMALYAIFVTWRMMGKNYD
                     AAVLAAGHCGFGLGATPTAIANMQAITERFGPSHMAFLVVPMVGAFFIDIVNALVIKL
                     YLMLPIFAG"
BASE COUNT     1040 a    1165 c    1231 g    1234 t
ORIGIN

  Query Match           46.0%;  Score 69;  DB 14;  Length 4670;
  Best Local Similarity  88.2%;  Pred. No. 2.4e-12;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1500 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1441

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db    1440 GGGGGCCCGGTACCCAGCTTTTGTT 1416


Search completed: May 29, 2000, 21:35:42
Job time: 36874 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:33 ; Search time 1446.77 Seconds
(without alignments)
25.940 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	69	46.0	259	1	Q87664	Mouse azoospermia
2	69	46.0	501	1	T04866	Nucleotide analogu
3	67	44.7	651	1	X20513	Polynucleotide seq
c 4	67	44.7	752	1	V31294	E. coli J96 pathog
c 5	67	44.7	1091	1	T75006	Human endogenous r
c 6	67	44.7	1122	1	T75005	Human endogenous r
c 7	67	44.7	3792	1	Q48463	Plasmid pG+host4 c
8	67	44.7	4226	1	T39485	Human steroidogene
c 9	67	44.7	5234	1	Q48464	Plasmid pG+host5 c
c 10	67	44.7	6722	1	Q48465	Plasmid pG+host6 c
c 11	67	44.7	10529	1	V09028	Maize fluory2 gene
12	67	44.7	12814	1	X24730	Swedish-FAD APP ta
13	67	44.7	15692	1	X24731	London-FAD APP tar
14	67	44.7	15692	1	X24732	Swedish/London-FAD
15	67	44.7	15701	1	X24733	Swedish-FAD APP713
c 16	66.8	44.5	545	1	V68808	Human endogenous r
17	65.8	43.9	6596	1	V57377	Maize female-prefe
c 18	65.4	43.6	2640	1	Q26664	bDAT. cDNA encodin
c 19	65.4	43.6	5356	1	T43794	Plasmid pRIPHAT (r
20	65	43.3	84	1	X02800	E. coli biotin DNA
c 21	65	43.3	3465	1	X02814	DE19731274 Seq ID
c 22	65	43.3	3481	1	X02815	DE19731274 Seq ID
c 23	65	43.3	3794	1	X02812	DE19731274 Seq ID
c 24	65	43.3	3810	1	X02813	DE19731274 Seq ID
c 25	63.6	42.4	545	1	T75010	Human endogenous r
26	62.4	41.6	685	1	V10190	Stealth virus nucl
27	62.4	41.6	685	1	V12003	Stealth virus plas
28	59	39.3	4164	1	T04575	Plasmid pAT-1 sequ
29	59	39.3	4164	1	V22271	pAT-1 (pSD544). DN
30	59	39.3	4164	1	V69740	Nucleotide sequenc
31	58.2	38.8	1949	1	X05602	Nucleotide sequenc
32	58	38.7	5178	1	T49876	pTet-Splice. Nucle
33	57.8	38.5	3198	1	T92702	Candida CaRho1 gen
34	57.8	38.5	3198	1	T92869	Candida CaRho1 gen
35	56.4	37.6	459	1	V57250	Clone #4 from muta
c 36	56	37.3	78	1	X02799	E. coli biotin DNA
c 37	56	37.3	2150	1	T91037	Yeast checkpoint c
c 38	56	37.3	2150	1	X01271	Yeast RAD17 coding
39	56	37.3	2973	1	V64254	Plasmid pKS varian
40	56	37.3	3956	1	V64258	Plasmid pPK13/14 D
41	56	37.3	4088	1	V64255	Plasmid pPK5/6 DNA
42	56	37.3	4102	1	V64257	Plasmid pPK9/10 DN
43	56	37.3	4583	1	V64256	Plasmid pPK7/8 DNA
44	54.4	36.3	1094	1	X29498	S. aureus MurF ORF
45	54	36.0	10504	1	V79503	DNA sequence of ex

ALIGNMENTS

RESULT 1
Q87664/c
ID Q87664 standard; DNA; 259 BP.
AC Q87664;
DT 06-MAR-1996 (first entry)
DE Mouse azoospermia factor (AZF) gene partial clone M3.2.
KW Azoospermia factor; AZF; male infertility; YRRM gene;
KW Y-chromosome; RNA recognition motif; ds.
OS Mus sp.
FH Key Location/Qualifiers
FT misc_difference 174
FT /*tag= a
FT /note= "base n at position 174 is not identified
FT in the specification"
FT misc_difference 197
FT /*tag= b
FT /note= "base n at position 197 is not identified
FT in the specification"
PN WO9511300-A2.
PD 27-APR-1995.
PF 24-OCT-1994; G02344.
PR 22-OCT-1993; GB-021857.
PR 07-JUL-1994; GB-013760.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;
DR WPI; 95-170221/22.
PT Nucleic acid encoding the human azoospermia factor, and probes and
PT antibodies specific for the sequence and encoded polypeptide - may
PT be used in the clinical diagnosis of male infertility
PS Disclosure; Fig 5; 40pp; English.
CC A mouse gene partial sequence (Q87664) showing homology to the human
CC azoospermia factor (AZF) YRMM gene was obtd. from mouse genomic phage
CC clone M3.2 isolated from a genomic library screened with MYK (YRRM1)
CC cDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm
CC of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also
CC obtd.
SQ Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;

Query Match 46.0%; Score 69; DB 1; Length 259;
Best Local Similarity 88.2%; Pred. No. 8.5e-16;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 53

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db      52 GGGGGCCCGGTACCCAGCTTTTGTT 28
```

RESULT 2
T04866
ID T04866 standard; DNA; 501 BP.
AC T04866;

```
DT   28-JAN-1996  (first entry)
DE   Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW   Nucleotide analogue; alkaline phosphatase; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .501
FT                   /*tag= a
FT                   /note= "std. IUPAC codes used"
PN   WO9524505-A.
PD   14-SEP-1995.
PF   07-MAR-1995; G00490.
PR   08-MAR-1994; EP-301636.
PA   (AMSH ) AMERSHAM INT PLC.
PI   Reeve MA,  Robinson PS;
DR   WPI; 95-328290/42.
PT   Modification of residual fluorescence labelled nucleotide analogues  - to
PT   prevent migration in electrophoretic sequencing gel and interference with
PT   base calling of DNA chains
PS   Example; Fig 1; 18pp; English.
CC   The invention concerns the modification of redisual fluorescence
CC   labelled nt analogues to prevent migration in electrophoretic
CC   sequencing gel and interference with base calling of DNA chains. The
CC   modification involves the use of a phosphatase enzyme to remove at
CC   least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC   dye-terminating kt. The template used was 1 microg. of M13 mp8. The
CC   primer was M13 Universal primer. Samples were recovered. Some were
CC   subjected to calf intestinal alkaline phospatase digestion. This
CC   includes T04866. T04867 was a control nt. which was not treated with
CC   alkaline phosphatase. The example shows that alkaline phosphatase
CC   treatment causes removal of the dye-terminator artefacts and allows
CC   for accurate base calling with the ABI analysis software.
SQ   Sequence   501 BP;    98 A;     136 C;    140 G;    121 T;

  Query Match             46.0%;  Score 69;  DB 1;  Length 501;
  Best Local Similarity   88.2%;  Pred. No. 1.1e-15;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 116

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     117 GGGGGCCCGGTACCCAGCTTTTGTT 141


RESULT   3
X20513
ID   X20513 standard; DNA; 651 BP.
AC   X20513;
DT   05-MAY-1999  (first entry)
DE   Polynucleotide sequence from the genome of Treponema pallidum.
KW   Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW   enzyme production; ds.
OS   Treponema pallidum.
PN   WO9859034-A2.
PD   30-DEC-1998.
PF   23-JUN-1998; U13041.
PR   24-JUN-1997; US-050667.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Fraser CM;
DR   WPI; 99-081273/07.
PT   New isolated Treponema pallidum nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of T. pallidum infections, particularly syphilis
PS   Claim 1; Page 257; 1150pp; English.
CC   X20500-21243 represent polynucleotide sequences from the genome of
CC   Treponema pallidum. The sequences can be used for detection,
CC   diagnosis, characterisation, prevention and therapy for T. pallidum
CC   infections, particularly syphilis. They can also be used for detecting
CC   diseases related to Borrelia infections in animals, and for the
CC   production of biosynthetic products such as enzymes.
SQ   Sequence   651 BP;    180 A;    167 C;    152 G;    149 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 137

Qy     111 GGGGGCC 117
           |||||||
Db     138 GGGGGCC 144


RESULT   4
V31294/c
ID   V31294 standard; DNA; 752 BP.
AC   V31294;
DT   01-OCT-1998  (first entry)
DE   E. coli J96 pathogenicity island contig #108.
KW   PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW   PAI V; pheV; vaccine; protective immune response; ds.
OS   Escherichia coli.
PN   WO9822575-A2.
PD   28-MAY-1998.
PF   21-NOV-1997; U21347.
PR   14-OCT-1997; US-061953.
PR   22-NOV-1996; US-031626.
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (UYWI-) UNIV WISCONSIN.
PI   Choi GH, Dillon PJ, Welch RA;
DR   WPI; 98-312461/27.
PT   New isolated uropathogenic E. coli nucleotide sequences - used to
PT   develop products for the detection of pathogenic E. coli and to
PT   elicit an immune response to pathogenic E. coli
PS   Claim 21; Page 221-222; 250pp; English.
CC   This sequence represents a E. coli strain J96 contig containing
CC   pathogenicity island (PAI) sequences, and represents a nucleic acid
CC   molecule of the invention. PAIs are large fragments of DNA which comprise
CC   pathogenicity determinants. The sequences of the invention are taken from
CC   PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC   on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC   approximately 94 min (at pheR) on the E. coli chromosome and is
CC   approximately 160 kb in size. Antibodies specific to the proteins encoded
CC   by the PAI open reading frames of the invention can be used in kits to
CC   detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC   a protective immune response in an animal to the uropathogenic E. coli
CC   strain J96.
SQ   Sequence   752 BP;    162 A;    213 C;    203 G;    172 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 752;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 13

Qy     111 GGGGGCC 117
           |||||||
Db      12 GGGGGCC 6


RESULT   5
T75006/c
ID   T75006 standard; DNA; 1091 BP.
AC   T75006;
DT   06-OCT-1997  (first entry)
DE   Human endogenous retroviral sequence 6.
KW   Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
```

OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
DR WPI; 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 31-32; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (T75003-10) were obtd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1091 BP; 79 A; 350 C; 97 G; 248 T;

Query Match 44.7%; Score 67; DB 1; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 78

Qy 111 GGGGGCC 117
|||||||
Db 77 GGGGGCC 71

RESULT 6
T75005/c
ID T75005 standard; DNA; 1122 BP.
AC T75005;
DT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 3.
KW Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
DR WPI; 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 30-31; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (T75003-10) were obtd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1122 BP; 260 A; 316 C; 185 G; 279 T;

Query Match 44.7%; Score 67; DB 1; Length 1122;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 49

Qy 111 GGGGGCC 117
|||||||
Db 48 GGGGGCC 42

RESULT 7
Q48463/c
ID Q48463 standard; DNA; 3792 BP.
AC Q48463;
DT 18-MAR-1994 (first entry)
DE Plasmid pG+host4 containing Ts replication system.
KW Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
OS Synthetic.
PN WO9318164-A.
PD 16-SEP-1993.
PF 12-MAR-1993; F00248.
PR 13-MAR-1992; FR-003034.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Gruss A, Maguin E;
DR WPI; 93-303478/38.
PT New bacterial plasmid contg. heat sensitive replication system -
PT and marker gene, opt. capable of chromosomal integration, used to
PT inactivate specific gene or introduce heterologous gene
PS Example 2; Fig 9; 73pp; French.
CC Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC site, T7 and T3 promoters and sequencing primer binding sites. The
CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463). It
CC is thermosensitive in all hosts tested, including E.coli, and must be
CC maintained at 28 deg.C.
SQ Sequence 3792 BP; 1249 A; 552 C; 742 G; 1249 T;

Query Match 44.7%; Score 67; DB 1; Length 3792;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy 111 GGGGGCC 117
|||||||
Db 3481 GGGGGCC 3475

RESULT 8
T39485
ID T39485 standard; DNA; 4226 BP.
AC T39485;
DT 21-MAY-1997 (first entry)
DE Human steroidogenesis acute regulatory protein genomic DNA.
KW Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW mutation; detection; prenatal; genetic defect; congenital; protein;
KW lipoid adrenal hyperplasia; treatment; prevention; gene;
KW replacement therapy; hypercholesterolaemia; ds.
OS Homo sapiens.
PN WO9629338-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03896.
PR 23-MAR-1995; US-410540.
PA (REGC) UNIV CALIFORNIA.
PA (UYPE-) UNIV PENNSYLVANIA.
PI Lin D, Miller WL, Strauss JF;
DR WPI; 96-443130/44.
PT Isolated human steroidogenesis acute regulatory protein gene - used

```
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
PS   Claim 1; Pages 23-25; 89pp; English.
CC   The present sequence encodes the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc abd pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
SQ   Sequence   4226 BP;    940 A;    1132 C;    1144 G;    984 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 4226;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4149 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4208

Qy     111 GGGGGCC 117
           |||||||
Db    4209 GGGGGCC 4215


RESULT   9
Q48464/c
ID   Q48464 standard; DNA; 5234 BP.
AC   Q48464;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host5 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    36. .1496
FT                   /*tag=  a
FT                   /standard_name= ORI
FT                   /note= "origin of replication from pBR322"
FT   misc_feature    2640. .4383
FT                   /*tag=  b
FT                   /note= "from pGK12 (derived from pWV01)"
FT   misc_feature    4384. .4786
FT                   /*tag=  c
FT                   /note= "from pUB110"
FT   misc_feature    4787. .5234
FT                   /*tag=  d
FT                   /note= "from pSK"
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Example 2; Fig 10; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the 1.4kb AvaI-AlwNI fragment of
CC   pBR322 (containing the origin of replication) was inserted into
CC   NsiI-cleaved pG+host4 to give pG+host5 (Q48464).The pBR322 ORI
CC   allows maintenance of the plasmid in E.coli at 37 deg.C. while the
CC   heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
CC   bacteria.
SQ   Sequence   5234 BP;    1569 A;    973 C;    1133 G;    1559 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy     111 GGGGGCC 117
           |||||||
Db    4923 GGGGGCC 4917


RESULT   10
Q48465/c
ID   Q48465 standard; DNA; 6722 BP.
AC   Q48465;
DT   18-MAR-1994  (first entry)
DE   Plasmid pG+host6 containing Ts replication system.
KW   Temperature sensitive replication; antibiotic resistance marker gene;
KW   site-specific recombination; chromosomal integration; inactivation;
KW   heterologous gene expression; thermosensitive plasmid; ds.
OS   Synthetic.
PN   WO9318164-A.
PD   16-SEP-1993.
PF   12-MAR-1993; F00248.
PR   13-MAR-1992; FR-003034.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PI   Gruss A,  Maguin E;
DR   WPI; 93-303478/38.
PT   New bacterial plasmid contg. heat sensitive replication system -
PT   and marker gene, opt. capable of chromosomal integration, used to
PT   inactivate specific gene or introduce heterologous gene
PS   Disclosure; Fig 11; 73pp; French.
CC   Plasmid pGK12 (Appl.Environ.Microbiol., 48; 726 (1984)) contg. two
CC   antibiotic resistance marker genes was subjected to mutagenesis with
CC   hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC   sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC   the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC   445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC   site, T7 and T3 promoters and sequencing primer binding sites. The
CC   resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).
CC   To facilitate cloning in E.coli, the AvaI-EcoRI fragment of pBR322
CC   (containing the origin of replication and the ampicillin resistance
CC   gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
CC   pG+host6 (Q48466).The pBR322 ORI allows maintenance of the plasmid in
CC   E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC   at 28 deg.C in gram-positive bacteria.
SQ   Sequence   6722 BP;    1961 A;    1278 C;    1503 G;    1980 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy     111 GGGGGCC 117
           |||||||
Db    6411 GGGGGCC 6405
```

```
RESULT  11
V09028/c
ID   V09028 standard; DNA; 10529 BP.
AC   V09028;
DT   20-JUL-1998  (first entry)
DE   Maize fluory2 gene (fl2).
KW   Floury2; fl2 gene; alpha-zein; signal peptide; transgenic plant;
KW   seed; essential amino acid; animal feed; maize; rice; wheat;
KW   barley; millet; sorghum; ds.
OS   Zea mays.
FH   Key             Location/Qualifiers
FT   misc_feature    1. .760
FT                   /*tag= a
FT                   /note= "vector-derived sequence"
FT   promoter        761. .3824
FT                   /*tag= b
FT                   /note= "Claim 14"
FT   CDS             3825. .4613
FT                   /*tag= c
FT   sig_peptide     3825. .3887
FT                   /*tag= d
FT   mat_peptide     3888. .4610
FT                   /*tag= e
PN   WO9802563-A1.
PD   22-JAN-1998.
PF   11-JUL-1997; U11723.
PR   17-JUL-1996; US-021833.
PA   (PION-) PIONEER HI-BRED INT INC.
PA   (UYAR-) UNIV ARIZONA STATE.
PI   Beach L, Coleman CE, Larkins BA;
DR   WPI; 98-110609/10.
DR   P-PSDB; W23977.
PT   Cereal plants containing trans-gene expressing fusion that includes
PT   signal peptide of the fl2 maize gene - and protein having high
PT   content of essential amino acids, producing feeds of improved
PT   nutritional value
PS   Claim 14; Fig 1A-H; 37pp; English.
CC   This is the nucleotide sequence of a clone of the fluory2 (fl2)
CC   gene of maize.  It codes for a 24-kDa alpha-zein protein (see
CC   W23977) that includes a 21-amino acid signal peptide (see W23976)
CC   which targets the alpha-zein to the lumen of the rough endoplasmic
CC   reticulum.  A claimed cereal plant contains a transgene comprising
CC   a first polynucleotide that encodes the fl2 signal peptide and a
CC   second polynucleotide that encodes an agronomically high-value
CC   protein.  Also new are seeds produced by the plants, the transgene
CC   itself, and a transgene that also includes the fl2 promoter.  The
CC   second polynucleotide preferably encodes a protein that has a high
CC   content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC   maize, wheat, rice, barley, millet or sorghum will have increased
CC   contents of these essential amino acids in their seeds (all
CC   claimed).
SQ   Sequence   10529 BP;    2845 A;    2484 C;    2262 G;    2938 T;


  Query Match            44.7%;  Score 67;  DB 1;  Length 10529;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 667

Qy    111 GGGGGCC 117
          |||||||
Db    666 GGGGGCC 660


RESULT  12
X24730
ID   X24730 standard; DNA; 12814 BP.
AC   X24730;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP targetting vector pMTI-2398.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2398;
KW   Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     1932. .2276
FT                   /*tag= a
FT                   /standard_name= Swedish-FAD APP
FT                   /note= "encodes W97997"
FT   mat_peptide     5369. .6160
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97997, W97998.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 96-103; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP targetting
CC   vector pMTI-2398.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5007,
CC   created using vector pMTI-2398, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced.  The targetted Swedish-FAD
CC   m/hAPP gene expressed m/hAPP protein at levels approaching those
CC   observed for mouse APP in brain.
SQ   Sequence   12814 BP;    1865 A;    1850 C;    1962 G;    1939 T;


  Query Match            44.7%;  Score 67;  DB 1;  Length 12814;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9888 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 9947

Qy    111 GGGGGCC 117
          |||||||
Db   9948 GGGGGCC 9954


RESULT  13
X24731
ID   X24731 standard; DNA; 15692 BP.
AC   X24731;
DT   21-JUN-1999  (first entry)
DE   London-FAD APP targetting vector pMTI-2453.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

```
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2453;
KW   London-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /standard_name= "Swedish-FAD"
FT                   /note= "encodes W97999"
FT   mutation        replace(4990,"")
FT                   /*tag= b
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= c
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W97999.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 104-113; 209pp; English.
CC   This is the nucleotide sequence of London-FAD APP targetting
CC   vector pMTI-2453. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5401,
CC   created using vector pMTI-2453, the London mutation. i.e.
CC   V(717)L, was also introduced.
SQ   Sequence   15692 BP;    2423 A;    2730 C;    2786 G;    2541 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12766 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12825

Qy     111 GGGGGCC 117
           |||||||
Db   12826 GGGGGCC 12832


RESULT  14
X24732
ID   X24732 standard; DNA; 15692 BP.
AC   X24732;
DT   21-JUN-1999  (first entry)
DE   Swedish/London-FAD APP targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2454;
KW   London-FAD APP; Swedish-FAD APP; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .5151
FT                   /*tag= a
FT                   /note= "encodes W98000"
FT   mutation        replace(4849,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4989,"")
FT                   /standard_name= "London-FAD"
FT   mat_peptide     8223. .9023
FT                   /*tag= b
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98000.
PT   Modification of target nucleic acids - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into.rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 114-123; 209pp; English.
CC   This is the nucleotide sequence of Swedish/London-FAD APP targetting
CC   vector pMTI-2454. The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step. The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells. In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18). Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain. In line ES5103,
CC   created using vector pMTI-2454, the London mutation. i.e.
CC   V(717)L, and the Swedish mutation, i.e. KM(670,671)NL, were
CC   introduced. The targetted Swedish/London-FAD m/hAPP gene expressed
CC   m/hAPP protein at levels approaching those observed for mouse APP
CC   in brain.
SQ   Sequence   15692 BP;    2423 A;    2731 C;    2783 G;    2542 T;

  Query Match             44.7%;  Score 67;  DB 1;  Length 15692;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12766 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12825

Qy     111 GGGGGCC 117
           |||||||
Db   12826 GGGGGCC 12832


RESULT  15
X24733
ID   X24733 standard; DNA; 15701 BP.
AC   X24733;
DT   21-JUN-1999  (first entry)
DE   Swedish-FAD APP713 targetting vector pMTI-2454.
KW   Amyloid precursor protein; APP; human; gene targetting;
```

```
KW   homologous recombination; transgenic mouse; transgenic animal;
KW   animal model; Alzheimer's disease; vector; pMTI-2455;
KW   Swedish-FAD APP713; mutation; ss.
OS   Homo sapiens.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   mat_peptide     4807. .4983
FT                   /*tag= a
FT                   /note= "encodes W98001"
FT   mutation        replace(4835,"")
FT                   /*tag= b
FT                   /standard_name= "Swedish-FAD"
FT   mutation        replace(4981,"")
FT                   /*tag= c
FT                   /standard_name= APP713stop
FT   mat_peptide     8232. .9032
FT                   /*tag= d
FT                   /note= "encodes W97998"
PN   WO9909150-A1.
PD   25-FEB-1999.
PF   18-AUG-1997; U14507.
PR   18-AUG-1997; WO-U14507.
PA   (FARB ) BAYER CORP.
PI   Wirak DO;
DR   WPI; 99-181029/15.
DR   P-PSDB; W97998, W98001.
PT   Modification of target nucleic acids  - by homologous recombination,
PT   used particularly for introducing a humanised amyloid precursor
PT   protein gene into rodents for producing models of Alzheimer's
PT   disease
PS   Example; Page 124-133; 209pp; English.
CC   This is the nucleotide sequence of Swedish-FAD APP713 targetting
CC   vector pMTI-2455.  The invention provides a novel gene targetting
CC   strategy that facilitates the introduction of one or more specific
CC   mutations into any gene in a single double reciprocal homologous
CC   recombination step.  The method has been used particularly for
CC   introducing a humanised amyloid precursor protein (APP) gene into
CC   rodents for producing animal models of Alzheimer's disease (AD).
CC   4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC   ES5401 and ES5403) have been created using the gene targetting
CC   technique applied to embryonic stem cells.  In each line, the mouse
CC   APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC   amino acid residues 666-770 of APP770 were encoded by human cDNA
CC   sequences instead of mouse genomic exons (exons 16-18).  Within
CC   these residues, only 3 amino acid differences exist between the
CC   mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC   Arg-684 to His.  The exon-cDNA fusion gene therefore encodes an APP
CC   containing a humanised beta-amyloid domain.  In line ES5215,
CC   created using vector pMTI-2455, the Swedish mutation. i.e.
CC   KM(670,681)NL, was also introduced, and also a stop codon (T to
CC   stop at position 714).
SQ   Sequence   15701 BP;    2470 A;    2675 C;    2845 G;    2497 T;


  Query Match             44.7%;  Score 67;  DB 1;  Length 15701;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12775 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 12834

Qy     111 GGGGGCC 117
           |||||||
Db   12835 GGGGGCC 12841


Search completed: May 29, 2000, 21:58:40
Job time: 38188 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:10 ; Search time 621.83 Seconds
(without alignments)
31.355 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn1_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn1_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn1_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn1_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn1_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	69.8	46.5	793	3	US-08-967-101-121	Sequence 121, App
c	2	69.8	46.5	793	4	US-08-592-541-121	Sequence 121, App
c	3	67	44.7	3792	4	US-08-992-334-1	Sequence 1, Appli
c	4	67	44.7	3792	5	US-08-302-752-1	Sequence 1, Appli
	5	67	44.7	4016	2	US-08-410-540-3	Sequence 3, Appli
c	6	67	44.7	5234	4	US-08-992-334-2	Sequence 2, Appli
c	7	67	44.7	5234	5	US-08-302-752-2	Sequence 2, Appli
c	8	67	44.7	6722	4	US-08-992-334-3	Sequence 3, Appli
c	9	67	44.7	6722	5	US-08-302-752-3	Sequence 3, Appli
	10	66.2	44.1	9318	3	US-08-793-610-6	Sequence 6, Appli
	11	62.4	41.6	685	2	US-08-463-115-56	Sequence 56, Appl
	12	62.4	41.6	685	2	US-08-465-388-56	Sequence 56, Appl
	13	59	39.3	4164	1	US-08-204-675-1	Sequence 1, Appli
	14	59	39.3	4164	3	US-08-660-754-1	Sequence 1, Appli
	15	59	39.3	4164	4	US-08-796-364-1	Sequence 1, Appli
	16	59	39.3	4164	6	PCT-US95-02520-1	Sequence 1, Appli
	17	58	38.7	5178	3	US-08-474-169-2	Sequence 2, Appli
	18	56	37.3	88	1	US-08-144-602B-15	Sequence 15, Appl
c	19	56	37.3	1023	1	US-08-198-446B-1	Sequence 1, Appli
c	20	56	37.3	1023	3	US-08-870-693-1	Sequence 1, Appli
c	21	56	37.3	2150	1	US-08-198-446B-10	Sequence 10, Appl
c	22	56	37.3	2150	3	US-08-870-693-10	Sequence 10, Appl
c	23	55	36.7	834	3	US-08-967-101-113	Sequence 113, App
c	24	55	36.7	834	4	US-08-592-541-113	Sequence 113, App
c	25	51.8	34.5	2308	1	US-08-325-071-62	Sequence 62, Appl
	26	51.2	34.1	3341	3	US-08-868-577-18	Sequence 18, Appl
c	27	50	33.3	780	3	US-08-967-101-116	Sequence 116, App
c	28	50	33.3	780	4	US-08-592-541-116	Sequence 116, App
	29	50	33.3	6505	3	US-08-793-610-5	Sequence 5, Appli
c	30	49.8	33.2	1200	1	US-08-011-398B-3	Sequence 3, Appli
c	31	49.8	33.2	1200	2	US-08-464-051-3	Sequence 3, Appli
c	32	49.8	33.2	1200	3	US-08-462-498-3	Sequence 3, Appli
c	33	49.8	33.2	6244	1	US-08-076-726-15	Sequence 15, Appl
c	34	49.8	33.2	6244	1	US-08-260-452-8	Sequence 8, Appli
c	35	49.8	33.2	6244	3	US-08-481-970-8	Sequence 8, Appli
c	36	49.8	33.2	6244	4	US-08-897-719-8	Sequence 8, Appli
	37	48	32.0	1612	1	US-08-343-733A-1	Sequence 1, Appli
	38	48	32.0	2764	4	US-08-465-971B-1	Sequence 1, Appli
c	39	47.2	31.5	2185	4	US-08-467-948A-3	Sequence 3, Appli
c	40	47	31.3	1997	4	US-08-667-809B-3	Sequence 3, Appli
c	41	47	31.3	2126	3	US-08-789-354-1	Sequence 1, Appli
	42	47	31.3	2605	3	US-08-680-395-4	Sequence 4, Appli
	43	47	31.3	5534	2	US-08-452-267-3	Sequence 3, Appli
c	44	46.6	31.1	1249	2	US-08-463-115-35	Sequence 35, Appl
c	45	46.6	31.1	1249	2	US-08-465-388-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT   1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-967-101-121


  Query Match             46.5%;  Score 69.8;  DB 3;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 4.5e-16;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   2
US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-592-541-121


  Query Match             46.5%;  Score 69.8;  DB 4;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 4.5e-16;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   3
US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
;    IMMEDIATE SOURCE:
;      CLONE:  pG+host4
US-08-992-334-1


  Query Match             44.7%;  Score 67;  DB 4;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy     111 GGGGGCC 117
           |||||||
Db    3481 GGGGGCC 3475


RESULT   4
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3792 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-1


  Query Match             44.7%;  Score 67;  DB 5;  Length 3792;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482

Qy     111 GGGGGCC 117
           |||||||
Db    3481 GGGGGCC 3475


RESULT   5
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F.
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,540
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-238/00US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4016 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1098..1283
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1620..1733
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2047..2174
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2267..2425
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2567..2751
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  2828..2921
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  3031..3765
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1433..1434
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2208..2209
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2781..2782
;      OTHER INFORMATION:  /note= "interruption of sequence data"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  2956..2957
;      OTHER INFORMATION:  /note= "interruption of sequence data"
US-08-410-540-3


  Query Match             44.7%;  Score 67;  DB 2;  Length 4016;
  Best Local Similarity   100.0%;  Pred. No. 8e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3939 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3998

Qy     111 GGGGGCC 117
           |||||||
Db    3999 GGGGGCC 4005


RESULT   6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
```

```
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-2


  Query Match             44.7%;  Score 67;  DB 4;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy     111 GGGGGCC 117
           |||||||
Db    4923 GGGGGCC 4917


RESULT   7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-2


  Query Match             44.7%;  Score 67;  DB 5;  Length 5234;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4983 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

Qy     111 GGGGGCC 117
           |||||||
Db    4923 GGGGGCC 4917


RESULT   8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
;  GENERAL INFORMATION:
;    APPLICANT:  Gruss, Alexandra
;    APPLICANT:  Maguin, Emmanuelle
;    TITLE OF INVENTION:  METHODS FOR USING A TEMPERATURE-SENSITIVE
;    TITLE OF INVENTION:  PLASMID
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christie Parker & Hale, LLP
;      STREET:  350 West Colorado Boulevard, Suite 500
;      CITY:  Pasadena
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/992,334
;      FILING DATE:  17-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/302,752
;      FILING DATE:  24-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00248
;      FILING DATE:  12-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR FR92/03034
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20958
;      REFERENCE/DOCKET NUMBER:  C93:31779
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (626) 795-9900
;      TELEFAX:  (626) 577-8800
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-992-334-3
```

```
  Query Match             44.7%;  Score 67;  DB 4;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy     111 GGGGGCC 117
           |||||||
Db    6411 GGGGGCC 6405


RESULT   9
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  THERMOSENSIBLE PLASMID
;    NUMBER OF SEQUENCES:  3
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/302,752
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 9203034
;      FILING DATE:  13-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO FR/93/00248
;      FILING DATE:  12-MAR-1993
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6722 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-302-752-3


  Query Match             44.7%;  Score 67;  DB 5;  Length 6722;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-15;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6471 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy     111 GGGGGCC 117
           |||||||
Db    6411 GGGGGCC 6405


RESULT  10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6


  Query Match             44.1%;  Score 66.2;  DB 3;  Length 9318;
  Best Local Similarity   95.8%;  Pred. No. 2.1e-14;
  Matches   68;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      47 TGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 106
           | | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5650 TAATGGGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 5709

Qy     107 GAGGGGGGGCC 117
           ||||||||||||
Db    5710 GAGGGGGGGCC 5720


RESULT  11
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
```

```
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,115
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:          Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5703221ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/301
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-463-115-56


  Query Match             41.6%;  Score 62.4;  DB 2;  Length 685;
  Best Local Similarity   94.0%;  Pred. No. 1.9e-13;
  Matches   63;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           |||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||||
Db     150 GTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGAGG 209

Qy     111 GGGGGCC 117
           ||||| |
Db     210 GGGGGNC 216


RESULT  12
US-08-465-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAM JOHN MARTIN
;    TITLE OF INVENTION:  ISOLATED STEALTH VIRUSES
;    TITLE OF INVENTION:  AND RELATED VACCINES
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,388
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:          Four
;      APPLICATION NUMBER:  08/157,811
;      FILING DATE:  No. 5753488ember 23, 1993
;      APPLICATION NUMBER:  07/887,502
;      FILING DATE:  May 22, 1992
;      APPLICATION NUMBER:  07/704,814
;      FILING DATE:  May 23, 1991
;      APPLICATION NUMBER:  07/763,039
;      FILING DATE:  September 20, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  213/300
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  685 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:
US-08-465-388-56


  Query Match             41.6%;  Score 62.4;  DB 2;  Length 685;
  Best Local Similarity   94.0%;  Pred. No. 1.9e-13;
  Matches   63;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           |||||||||||||| ||| |||||||||||||||| ||||||||||||||||||||||||
Db     150 GTGGATCCCCCGGCCTGNAGGAATTCGATATCAANCTTATCGATACCGTCGACCTCGAGG 209

Qy     111 GGGGGCC 117
           ||||| |
Db     210 GGGGGNC 216


RESULT  13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 5677170
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;.   APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,675
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-204-675-1


  Query Match             39.3%;  Score 59;  DB 1;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 6e-12;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  14
US-08-660-754-1
; Sequence 1, Application US/08660754
; Patent No. 5843772
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/660,754
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-660-754-1


  Query Match             39.3%;  Score 59;  DB 3;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 6e-12;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102


RESULT  15
US-08-796-364-1
; Sequence 1, Application US/08796364
; Patent No. 5968785
;  GENERAL INFORMATION:
;    APPLICANT:  Devine, Scott E.
;    APPLICANT:  Boeke, Jef D.
;    APPLICANT:  Braiterman, Lelita T.
;    TITLE OF INVENTION:  In Vitro Transposition of Artificial
;    TITLE OF INVENTION:  Transposons
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie, and Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/796,364
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,675
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  01107.45501
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202.508.9100
;      TELEFAX:  202.508.9299
;      TELEX:  197430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4164 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      CLONE:  pAT-1
US-08-796-364-1


  Query Match             39.3%;  Score 59;  DB 4;  Length 4164;
  Best Local Similarity   100.0%;  Pred. No. 6e-12;
```

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102
```

Search completed: May 29, 2000, 22:09:16
Job time: 38757 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:32 ; Search time 2276.24 Seconds
(without alignments)
294.022 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*
1: /cgn1_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn1_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn1_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn1_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn1_6/ptodata/1/pna/US081A_COMB.seq:*
6: /cgn1_6/ptodata/1/pna/US081B_COMB.seq:*
7: /cgn1_6/ptodata/1/pna/US081C_COMB.seq:*
8: /cgn1_6/ptodata/1/pna/US081D_COMB.seq:*
9: /cgn1_6/ptodata/1/pna/US082A_COMB.seq:*
10: /cgn1_6/ptodata/1/pna/US082B_COMB.seq:*
11: /cgn1_6/ptodata/1/pna/US082C_COMB.seq:*
12: /cgn1_6/ptodata/1/pna/US083A_COMB.seq:*
13: /cgn1_6/ptodata/1/pna/US083B_COMB.seq:*
14: /cgn1_6/ptodata/1/pna/US084A_COMB.seq:*
15: /cgn1_6/ptodata/1/pna/US084B_COMB.seq:*
16: /cgn1_6/ptodata/1/pna/US084C_COMB.seq:*
17: /cgn1_6/ptodata/1/pna/US085_COMB.seq:*
18: /cgn1_6/ptodata/1/pna/US086_COMB.seq:*
19: /cgn1_6/ptodata/1/pna/US087A_COMB.seq:*
20: /cgn1_6/ptodata/1/pna/US087B_COMB.seq:*
21: /cgn1_6/ptodata/1/pna/US087C_COMB.seq:*
22: /cgn1_6/ptodata/1/pna/US088A_COMB.seq:*
23: /cgn1_6/ptodata/1/pna/US088B_COMB.seq:*
24: /cgn1_6/ptodata/1/pna/US088C_COMB.seq:*
25: /cgn1_6/ptodata/1/pna/US088D_COMB.seq:*
26: /cgn1_6/ptodata/1/pna/US089A_COMB.seq:*
27: /cgn1_6/ptodata/1/pna/US089B_COMB.seq:*
28: /cgn1_6/ptodata/1/pna/US089C_COMB.seq:*
29: /cgn1_6/ptodata/1/pna/US089D_COMB.seq:*
30: /cgn1_6/ptodata/1/pna/US089E_COMB.seq:*
31: /cgn1_6/ptodata/1/pna/US089F_COMB.seq:*
32: /cgn1_6/ptodata/1/pna/US089G_COMB.seq:*
33: /cgn1_6/ptodata/1/pna/US089H_COMB.seq:*
34: /cgn1_6/ptodata/1/pna/US089I_COMB.seq:*
35: /cgn1_6/ptodata/1/pna/US090A_COMB.seq:*
36: /cgn1_6/ptodata/1/pna/US090B_COMB.seq:*
37: /cgn1_6/ptodata/1/pna/US090C_COMB.seq:*
38: /cgn1_6/ptodata/1/pna/US090D_COMB.seq:*
39: /cgn1_6/ptodata/1/pna/US090E_COMB.seq:*
40: /cgn1_6/ptodata/1/pna/US091A_COMB.seq:*
41: /cgn1_6/ptodata/1/pna/US091B_COMB.seq:*
42: /cgn1_6/ptodata/1/pna/US091C_COMB.seq:*
43: /cgn1_6/ptodata/1/pna/US092A_COMB.seq:*
44: /cgn1_6/ptodata/1/pna/US092B_COMB.seq:*
45: /cgn1_6/ptodata/1/pna/US092C_COMB.seq:*
46: /cgn1_6/ptodata/1/pna/US092D_COMB.seq:*
47: /cgn1_6/ptodata/1/pna/US092E_COMB.seq:*
48: /cgn1_6/ptodata/1/pna/US093A_COMB.seq:*
49: /cgn1_6/ptodata/1/pna/US093B_COMB.seq:*
50: /cgn1_6/ptodata/1/pna/US093C_COMB.seq:*
51: /cgn1_6/ptodata/1/pna/US093D_COMB.seq:*
52: /cgn1_6/ptodata/1/pna/US093E_COMB.seq:*
53: /cgn1_6/ptodata/1/pna/US093F_COMB.seq:*
54: /cgn1_6/ptodata/1/pna/US094A_COMB.seq:*
55: /cgn1_6/ptodata/1/pna/US094B_COMB.seq:*
56: /cgn1_6/ptodata/1/pna/US094C_COMB.seq:*
57: /cgn1_6/ptodata/1/pna/US095_COMB.seq:*
58: /cgn1_6/ptodata/1/pna/US6000_COMB.seq:*
59: /cgn1_6/ptodata/1/pna/US6001A_COMB.seq:*
60: /cgn1_6/ptodata/1/pna/US6001B_COMB.seq:*
61: /cgn1_6/ptodata/1/pna/US6002A_COMB.seq:*
62: /cgn1_6/ptodata/1/pna/US6002B_COMB.seq:*
63: /cgn1_6/ptodata/1/pna/US6002C_COMB.seq:*
64: /cgn1_6/ptodata/1/pna/US6003A_COMB.seq:*
65: /cgn1_6/ptodata/1/pna/US6003B_COMB.seq:*
66: /cgn1_6/ptodata/1/pna/US6004A_COMB.seq:*
67: /cgn1_6/ptodata/1/pna/US6004B_COMB.seq:*
68: /cgn1_6/ptodata/1/pna/US6004C_COMB.seq:*
69: /cgn1_6/ptodata/1/pna/US6005_COMB.seq:*
70: /cgn1_6/ptodata/1/pna/US6006_COMB.seq:*
71: /cgn1_6/ptodata/1/pna/US6007_COMB.seq:*
72: /cgn1_6/ptodata/1/pna/US6008A_COMB.seq:*
73: /cgn1_6/ptodata/1/pna/US6008B_COMB.seq:*
74: /cgn1_6/ptodata/1/pna/US6008C_COMB.seq:*
75: /cgn1_6/ptodata/1/pna/US6009A_COMB.seq:*
76: /cgn1_6/ptodata/1/pna/US6009B_COMB.seq:*
77: /cgn1_6/ptodata/1/pna/US6010A_COMB.seq:*
78: /cgn1_6/ptodata/1/pna/US6010B_COMB.seq:*
79: /cgn1_6/ptodata/1/pna/US6011A_COMB.seq:*
80: /cgn1_6/ptodata/1/pna/US6011B_COMB.seq:*
81: /cgn1_6/ptodata/1/pna/US6012A_COMB.seq:*
82: /cgn1_6/ptodata/1/pna/US6012B_COMB.seq:*
83: /cgn1_6/ptodata/1/pna/US6013A_COMB.seq:*
84: /cgn1_6/ptodata/1/pna/US6013B_COMB.seq:*
85: /cgn1_6/ptodata/1/pna/US6014A_COMB.seq:*
86: /cgn1_6/ptodata/1/pna/US6014B_COMB.seq:*
87: /cgn1_6/ptodata/1/pna/US6014C_COMB.seq:*
88: /cgn1_6/ptodata/1/pna/US6015_COMB.seq:*
89: /cgn1_6/ptodata/1/pna/US6016A_COMB.seq:*
90: /cgn1_6/ptodata/1/pna/US6016B_COMB.seq:*
91: /cgn1_6/ptodata/1/pna/US6016C_COMB.seq:*
92: /cgn1_6/ptodata/1/pna/US6017A_COMB.seq:*
93: /cgn1_6/ptodata/1/pna/US6017B_COMB.seq:*
94: /cgn1_6/ptodata/1/pna/US6017C_COMB.seq:*
95: /cgn1_6/ptodata/1/pna/US6018A_COMB.seq:*
96: /cgn1_6/ptodata/1/pna/US6018B_COMB.seq:*
97: /cgn1_6/ptodata/1/pna/US6018C_COMB.seq:*
98: /cgn1_6/ptodata/1/pna/US6019_COMB.seq:*
99: /cgn1_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
100: /cgn1_6/ptodata/1/pna/US06_NEW_COMB.seq:*
101: /cgn1_6/ptodata/1/pna/US07_NEW_COMB.seq:*
102: /cgn1_6/ptodata/1/pna/US08_NEW_COMB.seq:*
103: /cgn1_6/ptodata/1/pna/US09_NEW_COMB.seq:*
104: /cgn1_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	150	100.0	150	28	US-08-935-377-9	Sequence 9, Appli
2	138	92.0	149	28	US-08-935-377-8	Sequence 8, Appli

	3	136	90.7	148	28	US-08-935-377-7	Sequence 7, Appli
	4	130	86.7	145	28	US-08-935-377-6	Sequence 6, Appli
c	5	69.8	46.5	793	15	US-08-431-048A-121	Sequence 121, App
c	6	69.8	46.5	793	16	US-08-496-841-121	Sequence 121, App
c	7	69.8	46.5	793	17	US-08-509-359-121	Sequence 121, App
c	8	69.8	46.5	793	17	US-08-509-359B-121	Sequence 121, App
c	9	69.8	46.5	793	40	US-09-124-523-121	Sequence 121, App
c	10	69.8	46.5	793	40	US-09-124-698-121	Sequence 121, App
c	11	69.8	46.5	793	40	US-09-127-480-121	Sequence 121, App
	12	69	46.0	224	13	US-08-377-383-4	Sequence 4, Appli
	13	69	46.0	224	18	US-08-693-573-4	Sequence 4, Appli
c	14	67.8	45.2	1172	70	US-60-061-998-2	Sequence 2, Appli
c	15	67	44.7	147	38	US-09-072-433-31	Sequence 31, Appl
c	16	67	44.7	157	63	US-60-029-960-595	Sequence 595, App
	17	67	44.7	245	54	US-09-411-999-19130	Sequence 19130, A
	18	67	44.7	381	49	US-09-332-782-8570	Sequence 8570, Ap
	19	67	44.7	381	57	US-09-515-694-8570	Sequence 8570, Ap
	20	67	44.7	393	54	US-09-411-999-13627	Sequence 13627, A
	21	67	44.7	405	49	US-09-332-782-8523	Sequence 8523, Ap
	22	67	44.7	405	57	US-09-515-694-8523	Sequence 8523, Ap
	23	67	44.7	420	49	US-09-332-782-8595	Sequence 8595, Ap
	24	67	44.7	420	57	US-09-515-694-8595	Sequence 8595, Ap
	25	67	44.7	422	49	US-09-332-782-9309	Sequence 9309, Ap
	26	67	44.7	422	57	US-09-515-694-9309	Sequence 9309, Ap
	27	67	44.7	423	54	US-09-411-999-34199	Sequence 34199, A
	28	67	44.7	433	42	US-09-179-473-2334	Sequence 2334, Ap
	29	67	44.7	433	49	US-09-328-351-2334	Sequence 2334, Ap
	30	67	44.7	442	42	US-09-179-473-2495	Sequence 2495, Ap
	31	67	44.7	442	49	US-09-328-351-2495	Sequence 2495, Ap
	32	67	44.7	448	42	US-09-179-473-2255	Sequence 2255, Ap
	33	67	44.7	448	49	US-09-328-351-2255	Sequence 2255, Ap
	34	67	44.7	454	42	US-09-179-473-2100	Sequence 2100, Ap
	35	67	44.7	454	42	US-09-179-473-2360	Sequence 2360, Ap
	36	67	44.7	454	49	US-09-328-351-2100	Sequence 2100, Ap
	37	67	44.7	454	49	US-09-328-351-2360	Sequence 2360, Ap
c	38	67	44.7	752	32	US-08-976-259-108	Sequence 108, App
c	39	67	44.7	752	64	US-60-031-626-108	Sequence 108, App
c	40	67	44.7	827	82	US-60-126-265-996	Sequence 996, App
	41	67	44.7	858	40	US-09-123-912-91	Sequence 91, Appl
	42	67	44.7	858	43	US-09-221-107-91	Sequence 91, Appl
	43	67	44.7	858	46	US-09-285-479-91	Sequence 91, Appl
c	44	67	44.7	2958	37	US-09-040-961-6	Sequence 6, Appli
c	45	67	44.7	2958	41	US-09-130-115-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT   1
US-08-935-377-9
; Sequence 9, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  150 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-9


  Query Match           100.0%;  Score 150;  DB 28;  Length 150;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-42;
  Matches  150;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60

Qy      61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120

Qy     121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150
           ||||||||||||||||||||||||||||||
Db     121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150


RESULT   2
US-08-935-377-8
; Sequence 8, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-8
```

```
  Query Match             92.0%;  Score 138;  DB 28;  Length 149;
  Best Local Similarity   99.3%;  Pred. No. 1.3e-37;
  Matches  149;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGA-GTGGATCCCC 59

Qy     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 119

Qy    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150
          ||||||||||||||||||||||||||||||
Db    120 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 149


RESULT   3
US-08-935-377-7
; Sequence 7, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  148 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-7


  Query Match             90.7%;  Score 136;  DB 28;  Length 148;
  Best Local Similarity   98.7%;  Pred. No. 6.4e-37;
  Matches  148;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATG--GTGGATCCCC 58

Qy     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     59 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 118

Qy    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150
          ||||||||||||||||||||||||||||||
Db    119 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 148


RESULT   4
US-08-935-377-6
; Sequence 6, Application US/08935377
;  GENERAL INFORMATION:
;    APPLICANT:  Zauderer, Maurice
;    TITLE OF INVENTION:  T Cells Specific for Target Antigens and
;    TITLE OF INVENTION:  Vaccines Based Thereon
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C
;      STREET:  1100 New York Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/935,377
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1821.0010000/EKS/CMB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-935-377-6


  Query Match             86.7%;  Score 130;  DB 28;  Length 145;
  Best Local Similarity   96.7%;  Pred. No. 7.5e-35;
  Matches  145;  Conservative    0;  Mismatches    0;  Indels    5;  Gaps    1;

Qy      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGC-----CGTGGATCCCC 60
          ||||||||||||||||||||||||||||||||||||||||||||     ||||||||||
Db      1 GGCCAAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGC-----CGTGGATCCCC 55

Qy     61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     56 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCTAA 115

Qy    121 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 150
          ||||||||||||||||||||||||||||||
Db    116 CTAACTAATTTTGTTTTTGTGGGCCCGGCC 145


RESULT   5
US-08-431-048A-121/c
; Sequence 121, Application US/08431048A
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  132
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Administrator, TESTA, HURWITZ &
```

```
;      ADDRESSEE:  THIBEAULT, LLP
;      STREET:  125 High Street
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/431,048A
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38,349
;      REFERENCE/DOCKET NUMBER:  CAN-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-431-048A-121


  Query Match             46.5%;  Score 69.8;  DB 15;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   6
US-08-496-841-121/c
; Sequence 121, Application US/08496841
;  GENERAL INFORMATION:
;    APPLICANT:  St. George-Hyslop, Peter
;    APPLICANT:  Rommens, Johanna M.
;    APPLICANT:  Fraser, Paul E.
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/496,841
;      FILING DATE:  28-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-496-841-121


  Query Match             46.5%;  Score 69.8;  DB 16;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   7
US-08-509-359-121/c
; Sequence 121, Application US/08509359
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  160
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359
;      FILING DATE:  31-JUL-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Twomey, Michael J
;      REGISTRATION NUMBER:  38349
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7362
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359-121


  Query Match             46.5%;  Score 69.8;  DB 17;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           |||||||||||||||||||       |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   8
US-08-509-359B-121/c
; Sequence 121, Application US/08509359B
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
;      STREET:  600 South Avenue West
;      CITY:  Westfield
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07090
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word, Version 6.0c
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/509,359B
;      FILING DATE:  31-JUL-95
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Palisi, Thomas M.
;      REGISTRATION NUMBER:  36629
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 654-5000
;      TELEFAX:  (908) 654-7866
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-509-359B-121


  Query Match             46.5%;  Score 69.8;  DB 17;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           |||||||||||||||||||       |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT   9
US-09-124-523-121/c
; Sequence 121, Application US/09124523
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-523-121


  Query Match             46.5%;  Score 69.8;  DB 40;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           |||||||||||||||||||       |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  10
US-09-124-698-121/c
; Sequence 121, Application US/09124698
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,698
;      FILING DATE:
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-698-121


  Query Match             46.5%;  Score 69.8;  DB 40;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  11
US-09-127-480-121/c
; Sequence 121, Application US/09127480
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/127,480
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  121:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-127-480-121


  Query Match             46.5%;  Score 69.8;  DB 40;  Length 793;
  Best Local Similarity   82.5%;  Pred. No. 7.9e-14;
  Matches   80;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      39 GGCCGCCATGACGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
           ||  |||  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 GGAGGCCGCGGGGCGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71

Qy      99 TCGACCTCGAGGGGGGGCCTAACTAACTAATTTTGTT 135
           ||||||||||||||||||        |   |||||||
Db      70 TCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTT 34


RESULT  12
US-08-377-383-4
; Sequence 4, Application US/08377383
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru, FUKAMI,Tadashi, HIYOSHI,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of
;    TITLE OF INVENTION:  Nucleic Acid
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cushman Darby & Cushman
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskett, 3.5 inch, 720Kb storage
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1 or ASCII editors
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/377,383
;      FILING DATE:  25-JAN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6971/1994
;      FILING DATE:  26-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Glenn J. Perry
;      REGISTRATION NUMBER:  28458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-377-383-4


  Query Match             46.0%;  Score 69;  DB 13;  Length 224;
  Best Local Similarity   88.2%;  Pred. No. 1e-13;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     148 GGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT  13
```

```
US-08-693-573-4
; Sequence 4, Application US/08693573
;  GENERAL INFORMATION:
;    APPLICANT:  HOSOI,Shigeru,
;    APPLICANT:  FUKAMI,Tadashi,
;    APPLICANT:  KOJIMA,Makiko
;    TITLE OF INVENTION:  Method of Determining Base Sequence of                Nucleic
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY CUSHMAN
;      ADDRESSEE:  INTELLECTUAL PROPERTY GROUP OF
;      ADDRESSEE:  PILLSBURY MADISON & SUTRO, L.L.P.
;      STREET:  Ninth Floor, 1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 720Kb storage
;      COMPUTER:  IBM PC/XT/AT compatibles
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1 or ASCII editors
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,573
;      FILING DATE:  07-AUG-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/377,383
;      FILING DATE:  25-JAN-1995
;      APPLICATION NUMBER:  JP6971/1994
;      FILING DATE:  26-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Thomas G. Wiseman
;      REGISTRATION NUMBER:  35046
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other nucleic acid Synthetic DNA
US-08-693-573-4


  Query Match             46.0%;  Score 69;  DB 18;  Length 224;
  Best Local Similarity   88.2%;  Pred. No. 1e-13;
  Matches   75;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147

Qy     111 GGGGGCCTAACTAACTAATTTTGTT 135
           |||||||       |   |||||||
Db     148 GGGGGCCCGGTACCCAGCTTTTGTT 172


RESULT  14
US-60-061-998-2/c
; Sequence 2, Application US/60061998
;  GENERAL INFORMATION:
;    APPLICANT:  LAGACE, ROBERT E.
;    APPLICANT:  CORLEY, NEIL C.
;    APPLICANT:  RUSSO, FRANK D.
;    APPLICANT:  HANN, AMY L.
;    APPLICANT:  HEATH, JOE D.
;    APPLICANT:  FINNEY, GREGORY L.
;    TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
;    TITLE OF INVENTION:  GENOME, FRAGMENTS THEREOF, AND USES THEREOF
;    NUMBER OF SEQUENCES:  797
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/061,998
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PM-0006-2P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1172 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  SPN1c002
US-60-061-998-2


  Query Match             45.2%;  Score 67.8;  DB 70;  Length 1172;
  Best Local Similarity   75.7%;  Pred. No. 4.4e-13;
  Matches   84;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       5 AAAAATTGAAAAACTAGATCTATTTATTGCACGCGGCCGCCATGACGTGGATCCCCCGGG 64
             || ||   | ||   | || | |  | |   ||   | |  | |  | ||||||||||||
Db     111 AAGAAAAAATAATGGAAATATCTGAACTTGGCGTACCAGGTAAGGGGCGGATCCCCCGGG 52

Qy      65 CTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGG 115
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51 CTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGG 1


RESULT  15
US-09-072-433-31/c
; Sequence 31, Application US/09072433
;  GENERAL INFORMATION:
;    APPLICANT:  Coschigano, Peter W.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  BIOREMEDIATION
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,433
;      FILING DATE:  04-MAY-1998
;      CLASSIFICATION:  514
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Peter G.
;      REGISTRATION NUMBER:  32,837
;      REFERENCE/DOCKET NUMBER:  OHU-03344
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  147 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..147
US-09-072-433-31


  Query Match             44.7%;  Score 67;  DB 38;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-13;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     106 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 47

Qy     111 GGGGGCC 117
           |||||||
Db      46 GGGGGCC 40


Search completed: May 30, 2000, 09:48:32
Job time: 60182 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:16 ; Search time 2192.43 Seconds
(without alignments)
277.310 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCCAAAAATTGAAAAACTA.....TTGTTTTTGTGGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	70	46.7	682	96	AQ074693	AQ074693 CIT-HSP-2
c	2	68	45.3	608	96	AQ009167	AQ009167 CIT-HFP-2
c	3	67.4	44.9	528	84	B69688	B69688 CIT978SK-A-
	4	67	44.7	108	84	B54292	B54292 CIT-HSP-201
	5	67	44.7	141	44	AU037147	AU037147 AU037147
c	6	67	44.7	141	96	B95279	B95279 CIT-HSP-217
	7	67	44.7	144	84	B80294	B80294 CIT-HSP-204
c	8	67	44.7	169	96	AQ013255	AQ013255 CIT-HSP-2
	9	67	44.7	175	96	AQ041198	AQ041198 CIT-HSP-2
	10	67	44.7	259	20	T48593	T48593 ph6f4_19/1T
	11	67	44.7	275	27	AA052885	AA052885 T3551 Blo
c	12	67	44.7	386	84	B74162	B74162 CIT-HSP-202
c	13	67	44.7	462	96	AQ014512	AQ014512 CIT-HSP-2
c	14	67	44.7	519	84	B65534	B65534 CIT-HSP-202
	15	67	44.7	547	96	AQ012629	AQ012629 CIT-HSP-2
	16	67	44.7	552	96	AQ012216	AQ012216 CIT-HSP-2
	17	67	44.7	575	96	AQ012221	AQ012221 CIT-HSP-2
c	18	67	44.7	596	96	AQ009165	AQ009165 CIT-HFP-2
c	19	67	44.7	607	84	B64363	B64363 CIT-HSP-202
	20	67	44.7	632	84	B70299	B70299 CIT-HSP-202
c	21	67	44.7	638	96	AQ016737	AQ016737 CIT-HSP-2
	22	67	44.7	647	84	B68775	B68775 CIT-HSP-205
	23	67	44.7	666	84	B80296	B80296 CIT-HSP-204
	24	67	44.7	700	96	AQ009153	AQ009153 CIT-HSP-2
c	25	66.2	44.1	576	38	AA752172	AA752172 96BS0243
c	26	66	44.0	541	96	AQ074695	AQ074695 CIT-HSP-2
	27	65.4	43.6	513	96	AQ008437	AQ008437 CIT-HSP-2
c	28	65.4	43.6	541	96	AQ014519	AQ014519 CIT-HSP-2
c	29	62.8	41.9	134	84	B72088	B72088 CIT978SK-13
c	30	62.2	41.5	233	96	AQ012365	AQ012365 CIT-HSP-2
c	31	62	41.3	294	20	T26395	T26395 AB161A4F In
c	32	60.8	40.5	133	84	B26269	B26269 F10B5TF IGF
	33	60	40.0	100	43	AI239372	AI239372 SWOvAFCAP
c	34	57.8	38.5	145	84	B53695	B53695 CIT-HSP-201
c	35	57.8	38.5	630	96	AQ003238	AQ003238 RPCI11-19
c	36	57.4	38.3	603	38	AA752115	AA752115 96BS0066
	37	57.2	38.1	298	82	RPU68746	U68746 Riftia pach
	38	56.8	37.9	97	41	AI058116	AI058116 SWAMCAC15
c	39	56	37.3	144	84	B65391	B65391 CIT-HSP-202
	40	56	37.3	471	96	AQ024018	AQ024018 CpGR0359A
	41	56	37.3	499	96	AQ023564	AQ023564 CpGR0482A
	42	56	37.3	547	96	AQ023548	AQ023548 CpGR0051A
	43	56	37.3	626	96	AQ024077	AQ024077 CpGR0415A
c	44	55.4	36.9	502	96	AQ023674	AQ023674 CpGR0007B
c	45	55	36.7	149	96	AQ044029	AQ044029 CIT-HSP-2

ALIGNMENTS

RESULT 1
AQ074693/c
LOCUS AQ074693 682 bp DNA GSS 20-AUG-1998
DEFINITION CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23, genomic survey sequence.
ACCESSION AQ074693
VERSION AQ074693.1 GI:3436811
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 682)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2301L23"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"
BASE COUNT 156 a 187 c 187 g 151 t 1 others
ORIGIN

Query Match 46.7%; Score 70; DB 96; Length 682;
Best Local Similarity 93.6%; Pred. No. 1.4e-13;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 19

Qy     111 GGGGGCCTAACTAACTAA 128
           |||||||  | || | ||
Db      18 GGGGGCCCGAGTACCCAA 1
```

RESULT 2
AQ009167/c
LOCUS AQ009167 608 bp DNA GSS 27-JUN-1998
DEFINITION CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13, genomic survey sequence.
ACCESSION AQ009167
VERSION AQ009167.1 GI:3128599
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 608)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
CIT-HSP-2281N13.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .608

```
                     /organism="Homo sapiens"
                     /db_xref="GDB:7147360"
                     /db_xref="taxon:9606"
                     /clone="2281N13"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      141 a    164 c    166 g    137 t
ORIGIN


  Query Match             45.3%;  Score 68;  DB 96;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-13;
  Matches   68;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     111 GGGGGCCT 118
           ||||||||
Db      22 GGGGGCCT 15


RESULT   3
B69688/c
LOCUS       B69688         528 bp    DNA             GSS       18-JUN-1998
DEFINITION  CIT978SK-A-448E10.TVC CIT978SK Homo sapiens genomic clone A-448E10,
            genomic survey sequence.
ACCESSION   B69688
VERSION     B69688.1  GI:2708912
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
            Shizuya,H., Simon,M. and Venter,J.C.
  TITLE     Use of a human BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /db_xref="GDB:5299607"
                     /db_xref="taxon:9606"
                     /clone="A-448E10"
                     /clone_lib="CIT978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A"
BASE COUNT      118 a    138 c    147 g    125 t
ORIGIN


  Query Match             44.9%;  Score 67.4;  DB 84;  Length 528;
  Best Local Similarity   98.6%;  Pred. No. 1e-12;
  Matches   68;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 16

Qy     111 GGGGGCCTA 119
           ||||||| |
Db      15 GGGGGCCCA 7


RESULT   4
B54292
LOCUS       B54292         108 bp    DNA             GSS       20-JUN-1998
DEFINITION  CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
            genomic survey sequence.
ACCESSION   B54292
VERSION     B54292.1  GI:2608626
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7044003"
                     /db_xref="taxon:9606"
                     /clone="2017M1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       20 a     35 c     34 g     19 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 84;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 93

Qy     111 GGGGGCC 117
           |||||||
Db      94 GGGGGCC 100


RESULT   5
AU037147
```

```
LOCUS       AU037147      141 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSB532, mRNA sequence.
ACCESSION   AU037147
VERSION     AU037147.1  GI:3983900
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2153012.
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSB532"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      26 a     40 c     42 g     28 t      5 others
ORIGIN
```

```
  Query Match             44.7%;  Score 67;  DB 44;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 128

Qy     111 GGGGGCC 117
           |||||||
Db     129 GGGGGCC 135
```

```
RESULT   6
B95279/c
LOCUS       B95279        141 bp    DNA             GSS       26-JUN-1998
DEFINITION  CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
            genomic survey sequence.
ACCESSION   B95279
VERSION     B95279.1  GI:2977616
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 141)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21;
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .141
                     /organism="Homo sapiens"
                     /db_xref="GDB:7103988"
                     /db_xref="taxon:9606"
                     /clone="2172N2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      24 a     41 c     41 g     35 t
ORIGIN
```

```
  Query Match             44.7%;  Score 67;  DB 96;  Length 141;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 14

Qy     111 GGGGGCC 117
           |||||||
Db      13 GGGGGCC 7
```

```
RESULT   7
B80294
LOCUS       B80294        144 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
            genomic survey sequence.
ACCESSION   B80294
VERSION     B80294.1  GI:2867317
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2045D19.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="Homo sapiens"
                     /db_xref="GDB:7054558"
                     /db_xref="taxon:9606"
```

```
                     /clone="2045D19"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       34 a     43 c     41 g     26 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 84;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 129

Qy     111 GGGGGCC 117
           |||||||
Db     130 GGGGGCC 136


RESULT   8
AQ013255/c
LOCUS       AQ013255      169 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
            genomic survey sequence.
ACCESSION   AQ013255
VERSION     AQ013255.1  GI:3185820
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2299C22.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .169
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154018"
                     /db_xref="taxon:9606"
                     /clone="2299C22"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       32 a     47 c     48 g     42 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 96;  Length 169;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 42

Qy     111 GGGGGCC 117
           |||||||
Db      41 GGGGGCC 35


RESULT   9
AQ041198
LOCUS       AQ041198      175 bp    DNA             GSS       14-JUL-1998
DEFINITION  CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
            genomic survey sequence.
ACCESSION   AQ041198
VERSION     AQ041198.1  GI:3310469
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .175
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2335D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       43 a     50 c     47 g     35 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 96;  Length 175;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132

Qy     111 GGGGGCC 117
           |||||||
Db     133 GGGGGCC 139


RESULT  10
T48593
LOCUS       T48593        259 bp    mRNA            EST       02-FEB-1995
DEFINITION  ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph6f4_19/1TV, mRNA sequence.
```

```
ACCESSION   T48593
VERSION     T48593.1  GI:642793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L.,
            Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
            Sverdlov,E.D.
  TITLE     Outward Alu-primed hncDNA library
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Sverdlov ED
            Structure and Function of Human Genes
            Shemyakin Institute of Bioorganic Chemistry
            16/10 Miklukho-Maklaya, Moscow, 117871, Russia
            Tel: 70953306529
            Fax: 70953306538
            Email: sverd@humgen.siobc.msk.su.
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ph6f4_19/1TV"
                     /clone_lib="Outward Alu-primed hncDNA library"
                     /note="Vector: pGEM-3Z; Site_1: EcoR1; Site_2: BamH1; The
                     library was constructed as described in [Obradovic, D.,
                     Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
                     Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
                     Sverdlov, E.D. (1993) Bioorganicheskaya khimia, 20,
                     919-930]. This protocol is based on nested primer strategy
                     using Alu- specific primers (ALN3 and TC-65) that direct
                     the hncDNA synthesis outward of Alu repeats."
BASE COUNT       62 a     66 c     72 g     59 t
ORIGIN
```

```
  Query Match             44.7%;  Score 67;  DB 20;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 115

Qy     111 GGGGGCC 117
           |||||||
Db     116 GGGGGCC 122
```

```
RESULT  11
AA052885
LOCUS       AA052885      275 bp    mRNA            EST       13-SEP-1996
DEFINITION  T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
            brucei cDNA 5', mRNA sequence.
ACCESSION   AA052885
VERSION     AA052885.1  GI:1543898
KEYWORDS    EST.
SOURCE      Trypanosoma brucei brucei.
  ORGANISM  Trypanosoma brucei brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Osanya,A., Murphy,N.B. and Pelle,R.
  TITLE     Trypanosoma brucei brucei cDNAs
  JOURNAL   Unpublished (1996)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:637578.
            Contact: Osanya A
            Unit 3
            International Livestock Research Institute
            Box 30709, Nairobi, Kenya
            Tel: 254 2 630 743
            Fax: 254 2 631 499
            Email: a.osanya@cgnet.com
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Trypanosoma brucei brucei"
                     /strain="Clone A4"
                     /db_xref="taxon:5702"
                     /clone_lib="Bloodstream form of serodeme ILTat1.1"
                     /note="cDNAs were generated from poly (A+) enriched mRNA
                     prepared from different developmental stages of T.b.brucei
                     by reverse transcription followed by PCR amplification
                     using mini-exon and oligo(dT) primers.  The cDNA generated
                     were utilized in RADES-PCR coupled with differential
                     hybridisations to identify differentially expressed mRNA
                     transcripts. The products which showed to be
                     differentially expressed were cloned pGEM -T vector.
                     These differentially expressed mRNA transcripts were (are
                     being) sequenced to generate differentially expressed
                     sequence tags."
BASE COUNT       61 a     84 c     74 g     56 t
ORIGIN
```

```
  Query Match             44.7%;  Score 67;  DB 27;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 108

Qy     111 GGGGGCC 117
           |||||||
Db     109 GGGGGCC 115
```

```
RESULT  12
B74162/c
LOCUS       B74162        386 bp    DNA             GSS       24-OCT-1998
DEFINITION  CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
            genomic survey sequence.
ACCESSION   B74162
VERSION     B74162.1  GI:2769849
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 386)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2028N1.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .386
                     /organism="Homo sapiens"
                     /db_xref="GDB:7048251"
```

```
                     /db_xref="taxon:9606"
                     /clone="2028N1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       81 a     98 c    109 g     98 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 84;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     111 GGGGGCC 117
           |||||||
Db      22 GGGGGCC 16


RESULT  13
AQ014512/c
LOCUS       AQ014512      462 bp    DNA             GSS       06-JUN-1998
DEFINITION  CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
            genomic survey sequence.
ACCESSION   AQ014512
VERSION     AQ014512.1  GI:3184839
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154425"
                     /db_xref="taxon:9606"
                     /clone="2300D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       89 a    127 c    131 g    115 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 96;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 40

Qy     111 GGGGGCC 117
           |||||||
Db      39 GGGGGCC 33


RESULT  14
B65534/c
LOCUS       B65534        519 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
            genomic survey sequence.
ACCESSION   B65534
VERSION     B65534.1  GI:2639512
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /db_xref="GDB:7045730"
                     /db_xref="taxon:9606"
                     /clone="2022D24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      110 a    137 c    148 g    124 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 84;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23

Qy     111 GGGGGCC 117
           |||||||
Db      22 GGGGGCC 16


RESULT  15
AQ012629
LOCUS       AQ012629      547 bp    DNA             GSS       06-JUN-1998
```

```
DEFINITION  CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
            genomic survey sequence.
ACCESSION   AQ012629
VERSION     AQ012629.1  GI:3185194
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are availabe from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154069"
                     /db_xref="taxon:9606"
                     /clone="2299F1"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      154 a    139 c    125 g    129 t
ORIGIN


  Query Match             44.7%;  Score 67;  DB 96;  Length 547;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      51 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 134

Qy     111 GGGGGCC 117
           |||||||
Db     135 GGGGGCC 141


Search completed: May 29, 2000, 21:13:17
Job time: 36516 sec
```